JOURNAL OF ATMOSPHERIC AND TERRESTRIAL PHYSICS

EDITORIAL ADVISORY BOARD

H. O. G. ALFVÉN–STOCKHOLM · **J. BARTELS**–GÖTTINGEN · **W. J. G. BEYNON**–ABERYSTWYTH · **P. M. S. BLACKETT**–LONDON · **H. G. BOOKER**–ITHACA, N.Y. **E. G. BOWEN**–SYDNEY · **SIR DAVID BRUNT**–LONDON · **R. BUREAU**–BAGNEUX **S. CHAPMAN**–ALASKA · **J. EGEDAL**–COPENHAGEN · **V. F. HESS**–NEW YORK · **E. O. HULBURT**–WASHINGTON, D.C. · **H. ISRAEL**–AACHEN · **D. K. MENZEL**–CAMBRIDGE, MASS. · **S. K. MITRA**–CALCUTTA · **M. NICOLET**–UCCLE · **E. H. NORINDER**–UPSALA · **B. van der POL**–GENEVA · **N. V. PUSHKOV**–MOSCOW · **J. A. RATCLIFFE**–CAMBRIDGE · **O. E. H. RYDBECK**–GOTHENBURG · **J. C. W. SCOTT**–OTTAWA **SIR GEORGE SIMPSON**–LONDON · **H. SOLBERG**–OSLO · **J. VELDKAMP**–DE BILT **M. WALDMEIER**–ZÜRICH · **J. ZENNECK**–ALTHEGNENBERG

EDITOR-IN-CHIEF

SIR EDWARD APPLETON, F.R.S.

EDINBURGH

VOL. 14 1959

VOLUME 14

1959

PERGAMON PRESS

LONDON · NEW YORK · PARIS · LOS ANGELES

J. Atmosph. Terr. Phys.

VOL
14
195

CONTENTS

Nos. 1/2

W. R. French, Jr. and R. L. Chasson: Atmospheric effects on the hard component of cosmic radiation near sea level 1

A. A. Weiss: The limitations of narrow-beam radio equipments in the detection of weak meteor showers 19

R. G. Rastogi: Geomagnetic influence on the *F1*- and *F2*-regions of the ionosphere—effect of solar activity 31

J. W. King: Ionospheric self-demodulation and self-distortion of radio waves . . 41

J. E. Titheridge: Ray paths in the ionosphere. Approximate calculations in the presence of the earth's magnetic field 50

G. C. Reid and C. Collins: Observations of abnormal VHF radio wave absorption at medium and high altitudes 63

L. A. Manning: Oblique echoes from over-dense meteor trails 82

B. Ramachandra Rao and E. Bhagiratha Rao: Study of horizontal drifts in the *F1*- and *F2*-regions of the ionosphere at Waltair (17° 43′N, 83° 18′E, mag. lat. 9° 30′N) . . 94

L. Harang and J. Tröim: Determination of the angle of arrival of auroral echoes . 107

W. N. Abbott and S. Chapman: On the aurora of 19th August 1950, photographed from Greece 111

L. Thomas: Some measurements of horizontal movements in region-*F2* using widely spaced observing stations 123

W. J. G. Beynon and G. M. Brown: Geomagnetic distortion of region-*E* 138

S. K. Runcorn: On the hypothesis that the mean geomagnetic field for parts of geological time has been that of a geocentric axial multipole 167

Research Notes:

W. R. Piggott: A rapid method for obtaining accurate virtual heights from an ionogram 175

A. M. Arthurs: Cross-sections for *K*-shell ionization of atomic nitrogen and oxygen by electrons and protons 177

W. Dieminger, H. G. Möller and G. Rose: Further results of sweep-frequency oblique-incidence pulse transmissions 179

K-indices of geomagnetic activity for October and November 1958 181

Papers to be published in future issues 183

Nos. 3/4

G. McK. Allcock: The electron density distribution in the outer ionosphere derived from whistler data 185

R. Montalbetti: Photoelectric measurements of hydrogen emissions in aurorae and airglow 200

C. M. Minnis and G. H. Bazzard: Some indices of solar activity based on ionospheric and radio noise measurements 213

C. H. Cummack: Chapman behaviour in the lower ionosphere 229

G. S. Aitchison and K. Weekes: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—I. The theory of the analysis . . 236

G. S. Aitchison, J. H. Thomson and K. Weekes: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—II. Experimental procedure and results 244

C. A. Ventrice and E. R. Schmerling: Coefficients for the rapid reduction of *h′–f* records to *N–h* profiles without computing aids 249

F. Hepburn: Interpretation of smooth type atmospheric wave forms 262

VOL. 14 1959

CONTENTS

Nos. 1/2

W. R. French, Jr. and R. L. Chasson: Atmospheric effects on the hard component of cosmic radiation near sea level 1

A. A. Weiss: The limitations of narrow-beam radio equipments in the detection of weak meteor showers 19

R. G. Rastogi: Geomagnetic influence on the *F1*- and *F2*-regions of the ionosphere—effect of solar activity 31

J. W. King: Ionospheric self-demodulation and self-distortion of radio waves . . 41

J. E. Titheridge: Ray paths in the ionosphere. Approximate calculations in the presence of the earth's magnetic field 50

G. C. Reid and C. Collins: Observations of abnormal VHF radio wave absorption at medium and high altitudes 63

L. A. Manning: Oblique echoes from over-dense meteor trails 82

B. Ramachandra Rao and E. Bhagiratha Rao: Study of horizontal drifts in the *F1*- and *F2*-regions of the ionosphere at Waltair (17° 43′N, 83° 18′E, mag. lat. 9° 30′N) . . 94

L. Harang and J. Tröim: Determination of the angle of arrival of auroral echoes . 107

W. N. Abbott and S. Chapman: On the aurora of 19th August 1950, photographed from Greece 111

L. Thomas: Some measurements of horizontal movements in region-*F2* using widely spaced observing stations 123

W. J. G. Beynon and G. M. Brown: Geomagnetic distortion of region-*E* 138

S. K. Runcorn: On the hypothesis that the mean geomagnetic field for parts of geological time has been that of a geocentric axial multipole 167

Research Notes:

W. R. Piggott: A rapid method for obtaining accurate virtual heights from an ionogram 175

A. M. Arthurs: Cross-sections for *K*-shell ionization of atomic nitrogen and oxygen by electrons and protons 177

W. Dieminger, H. G. Möller and G. Rose: Further results of sweep-frequency oblique-incidence pulse transmissions 179

K-indices of geomagnetic activity for October and November 1958 181

Papers to be published in future issues 183

Nos. 3/4

G. McK. Allcock: The electron density distribution in the outer ionosphere derived from whistler data 185

R. Montalbetti: Photoelectric measurements of hydrogen emissions in aurorae and airglow 200

C. M. Minnis and G. H. Bazzard: Some indices of solar activity based on ionospheric and radio noise measurements 213

C. H. Cummack: Chapman behaviour in the lower ionosphere 229

G. S. Aitchison and K. Weekes: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—I. The theory of the analysis . . 236

G. S. Aitchison, J. H. Thomson and K. Weekes: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—II. Experimental procedure and results 244

C. A. Ventrice and E. R. Schmerling: Coefficients for the rapid reduction of *h′–f* records to *N–h* profiles without computing aids 249

F. Hepburn: Interpretation of smooth type atmospheric wave forms 262

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 1 to 18. Pergamon Press Ltd. Printed in Northern Ireland

Atmospheric effects on the hard component of cosmic radiation near sea level

W. R. French, Jr.* and R. L. Chasson
University of Nebraska, Lincoln, Nebraska

(*Received* 12 *May* 1958)

Abstract—Atmospheric effects on the hard component of cosmic radiation were studied at Lincoln, Nebraska (51° N geomagnetic latitude; 350 m altitude) with narrow- and wide-angle Geiger-counter telescopes (20 cm Pb absorber). Intensity measurements were made from 1 Sept. 1953 through 16 Feb. 1956, and linear regression analyses were carried out for several possible combinations of atmospheric variables and the cosmic-ray intensity. A detailed comparison is made of the results obtained from triple-regression calculations based upon the 50 mb, 100 mb and 200 mb Duperier models and the Olbert model. It is found that the 50 mb Duperier model is superior to all others tested, but the extent to which the intensity can be corrected by statistical methods is limited at present by the accuracy of the radiosonde data. Experimental results are in good general agreement with those of other investigators and agree well with the theoretical predictions of Trefall with regard to the temperature effect. They are in poor agreement with the predictions of Trefall on the barometric effect, however, and do not substantiate the theoretical conclusions of Olbert. The smoothing of the intensity curves seems to be practically indifferent to the terms employed in the regression equation if it involves three or more atmospheric variables.

1. Introduction

It is to be expected ideally that intensity changes of the primary cosmic rays are accompanied by synchronous changes of the secondary intensity observed at ground stations, where long-term monitor observations are possible. The ideal correspondence between primary and secondary fluctuations is considerably overshadowed, however, by large effects due to changing conditions of the atmosphere. Until a reliable system of intensity corrections related to the atmospheric effects is devised, therefore, one cannot hope to perform a precise extrapolation from observed secondary changes to the parent primary variations. This problem is not of major concern in the study of the low-energy nucleonic component, where decay effects do not enter, but it is a serious one in the study of the behaviour of the higher-energy fraction of the radiation, which includes the non-solar component. Its solution is of especial current importance with regard to the optimum quantitative utilization of the worldwide synoptic cosmic-ray monitor data that are being collected during the International Geophysical Year.

Two main lines of endeavour have been pursued by various investigators toward the solution of this problem, with the ultimate objective being the provision of a reliable and reasonable facile correction system. One approach has been purely statistical, with the objective of finding a set of partial regression coefficients relating μ-meson intensity changes to fluctuations of such atmospheric parameters as barometric pressure, height of various isobaric levels and temperatures of various strata (Duperier, 1949). The results have been, in turn, related with only partial success to known meson production, decay and absorption

* Now at Nebraska Wesleyan University, Lincoln, Nebraska.

phenomena (TREFALL, 1955a,b,c; 1957). The consistency of the results improved markedly as radiosonde data became available for higher regions of the atmosphere.

The second approach was phenomenological, where continuous production and decay of mesons were taken into account along with the ionization loss by the particles traversing the atmosphere (OLBERT, 1953; DORMAN *et al.*, 1953). The analytical work of OLBERT gave results in the form of a regression equation with theoretical partial regression coefficients for the atmospheric parameters, including a production-level coefficient. The work of DORMAN and his colleagues took into account the mass absorption and the air–mass distribution and thus eliminated the obviously idealized (but practically convenient) concept of a meson production "level" and simplified temperature effect. Application of the latter technique to correction procedures for observed intensity gave acceptable results (summarized in DORMAN and FEINBERG, 1955). The applicability of a successful phenomenological scheme such as this is reduced, however, when large quantities of data must be corrected on a routine basis; one is compelled to seek, therefore, a more facile scheme for routine analysis. It would seem that a consistent statistical system is operationally acceptable even though it may not be entirely satisfactory physically. Knowledge of the details of the behaviour of the atmospheric cascade is ultimately required, however, if small fluctuations such as a solar diurnal wave are to be analysed with high precision.

It is the primary purpose of this work to use greatly improved statistics to compare various statistical atmospheric "models" for cosmic-ray corrections. Second, it is intended to indicate the relationship of the accuracy of available radiosonde data to the ultimate accuracy of corrections obtained statistically. Third, a comparison will be made between simultaneous results for the radiation received by both wide- and narrow-angle counter telescopes.

VOL
14
195

2. EXPERIMENTAL ARRANGEMENT

The experiment consisted of two associated triple-coincidence Geiger-counter telescopes having narrow and wide opening angle, namely:

telescope A $49^\circ \times 80^\circ$
telescope B $135^\circ \times 80^\circ$

The telescope axes were vertical, with the common opening angle in the geographical east–west direction. The first four counters per tray constituted telescope A. Pulses from the four counters were split and then mixed electronically with those from sixteen adjoining counters in the tray, making a total of twenty counters in each tray of telescope B.

The Geiger-counter tubes used were of the external-cathode type filled with ethyl alcohol and argon. All counters were operated at the same voltage from a common regulated high voltage supply. A total thickness of 20 cm of lead absorber, distributed into two equal parts and placed between pairs of trays, served to define the lower momentum limit for the μ-mesons accepted (approximately 370 MeV/c). Assuming a $p^{-1.8}$ integral momentum spectrum, the average momentum of the

incident μ-mesons able to traverse the telescope vertically was approximately 460 MeV/c.

This dual scheme provided a common environment for both telescopes and made more efficient use of Geiger tubes. The approximate mean triple-coincidence rates were:

telescope A — 3000/hr (scaled by 4)

telescope B — 35,000/hr (scaled by 16)

with effective tray areas of 390 cm^2 and 1950 cm^2, respectively. The total roof thickness over the apparatus was estimated to be about 20 g/cm^2. The laboratory temperature was maintained at 23° $\pm$ 3°C.

The cosmic-ray laboratory is at Lincoln, Nebraska, which is at 51° N geomagnetic latitude (after McNish, 1936), and radiosonde data were taken from the WBAN 33 Omaha, Nebraska, reports supplied by the United States Weather Bureau. Omaha is approximately 80 km east of Lincoln. Barometric data were obtained from the Weather Bureau Station in Lincoln. The laboratory elevation is approximately 350 m.

3. Preliminary Results

The initial procedure in the calculation of results was to make a series of preliminary regression analyses using selected intensity data and various combinations of those atmospheric variables suspected of being intimately related to the μ-meson intensity (Chasson and French, 1955). The following variables were used initially:

I_A = narrow-angle intensity
I_B = wide-angle intensity
H_5 = height of the 50 mb level
H_1 = height of the 100 mb level
T_{12} = mean temperature between 100–200 mb levels (Duperier)
T_{52} = mean temperature between 50–200 mb levels (Duperier)
T_t = mean temperature of the troposphere (Olbert)

The Olbert temperature was computed by taking the arithmetic mean of the radiosonde temperatures over 50 mb intervals from the 950 mb to the 200 mb level; it was used only in computations involving the narrow-angle intensity data.

Selected for this preliminary analysis were certain days from the 5 month period November 1953, through March 1954, which was at the minimum of the solar-activity cycle. Data for these days met the following criteria:

(a) No geomagnetic or solar disturbances reported.

(b) The 0300 UT radiosonde reached at least 50 mb.

(c) Reliable intensity data were present for both narrow- and wide-angle telescopes.

A total of seventy-nine datum periods were found to meet these requirements. Intensity data for the 3 hr period bracketing the time of the radiosonde flights were used. Although this limitation reduced the overall counting statistics, it did not seem reasonable to take intensity data for the whole or even half-day to use with upper altitude data covering only a brief portion of that time (Chasson, 1954). The restriction to using only the 0300 UT radiosonde flight for this

calculation was arbitrary, but both radiosonde flights were used to make the final calculations presented in Section 4.

In all of the calculations the following linear regression equation was used:

$$\frac{\delta I}{\bar{I}} = \sum_i b_i \, \delta X_i \tag{1}$$

where δI is the departure of the cosmic-ray intensity from its mean value, $\bar{I}$; b_i is the regression coefficient expressed as percentage change per unit departure of the particular atmospheric variable X_i from its mean value. Table 1 presents a

Table 1. Results of preliminary analyses, using seventy-nine datum periods selected from the period November 1953—March 1954. Regression coefficients are shown together with standard errors. Refer to text (Section 3) for definition of group and variable symbols. R is the multiple regression coefficient.

Prelim. anal. group	Included variables	Partial barometric coefficient (%/cm Hg)	Partial height coefficient (%/km)	Duperier partial temperature coefficient (%/°C)	Olbert partial temperature coefficient (%/°C)	R
1A	$I_A B$	$-1{\cdot}30 \pm 0{\cdot}19$	—	—	—	0·62
2A	$I_A BH_1$	$-1{\cdot}32 \pm 0{\cdot}18$	$-4{\cdot}63 \pm 1{\cdot}44$	—	—	0·68
3A	$I_A BH_1 T_{12}$	$-1{\cdot}28 \pm 0{\cdot}17$	$-3{\cdot}36 \pm 1{\cdot}41$	$+0{\cdot}093 \pm 0{\cdot}029$	—	0·72
4A	$I_A BH_1 T_{12} T_t$	$-1{\cdot}73 \pm 0{\cdot}56$	$+0{\cdot}53 \pm 4{\cdot}85$	$+0{\cdot}0069 \pm 0{\cdot}106$	$-0{\cdot}156 \pm 0{\cdot}186$	0·73
5A	$I_A BH_1 T_t$	$-1{\cdot}76 \pm 0{\cdot}21$	$+0{\cdot}84 \pm 2{\cdot}11$	—	$-0{\cdot}168 \pm 0{\cdot}050$	0·73
6A	$I_A BH_5 T_{52}$	$-1{\cdot}26 \pm 0{\cdot}17$	$-2{\cdot}56 \pm 1{\cdot}49$	$+0{\cdot}154 \pm 0{\cdot}030$	—	0·73
1B	$I_B B$	$-1{\cdot}48 \pm 0{\cdot}11$	—	—	—	0·83
2B	$I_B BH_1$	$-1{\cdot}49 \pm 0{\cdot}10$	$-3{\cdot}58 \pm 0{\cdot}81$	—	—	0·87
3B	$I_B BH_1 T_{12}$	$-1{\cdot}46 \pm 0{\cdot}08$	$-2{\cdot}47 \pm 0{\cdot}71$	$+0{\cdot}080 \pm 0{\cdot}014$	—	0·91
4B	$I_B BH_5 T_{52}$	$-1{\cdot}44 \pm 0{\cdot}09$	$-1{\cdot}19 \pm 0{\cdot}76$	$+0{\cdot}127 \pm 0{\cdot}017$	—	0·91

summary of the regression coefficients and multiple correlation coefficients for various combinations of included variables, starting with the simple barometric coefficient, which assumes all variations to be due to changes in station barometric pressure, and going to such complicated cases as 4A, in which the regression equation has four terms (a five-variate multiple-regression analysis). No detailed discussion of these results will be made here since the results of more extensive calculations will be found in Section 4, and a discussion of all results is presented in Section 5.

The effects of "correcting" the intensity data according to the regression coefficients presented in Table 1 are shown in Figs. 1 and 2. It will be noted that the smoothing of the intensity curves is practically indifferent to the terms employed in the regression equation as long as it involves three or more atmospheric parameters.

4. FINAL RESULTS

Calculations were guided by the preliminary results reported above, and the work of DUPERIER (1949, 1951, 1956), OLBERT (1953), and TREFALL (1955a,b,c).

VOL.
14
1959

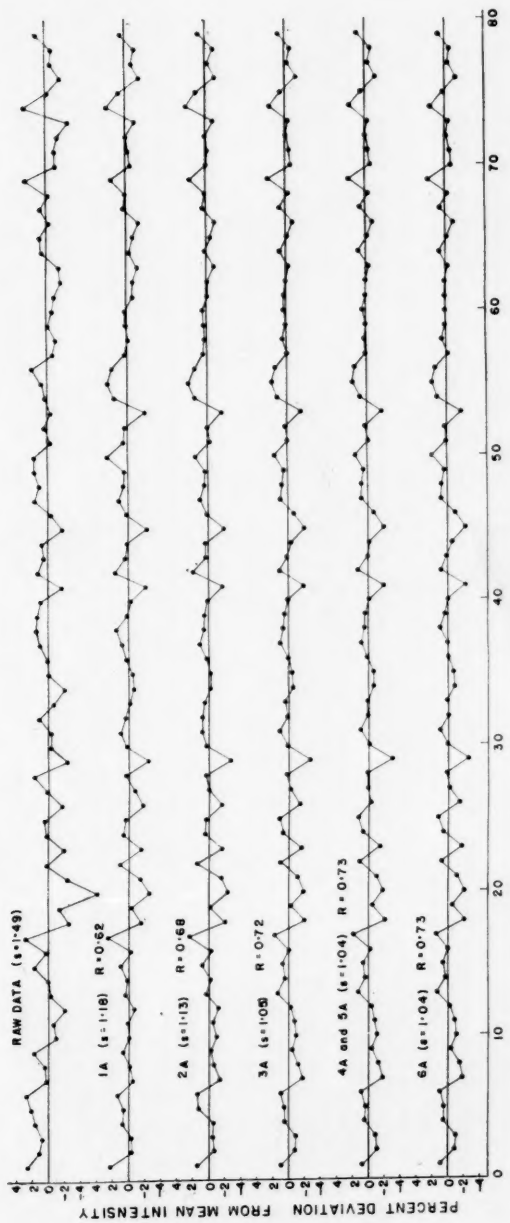


Fig. 1. Results of correcting narrow-angle cosmic-ray intensity according to regression coefficients found from preliminary analyses. (See Table 1 for identification of symbols.) S is the standard deviation of the points plotted for each group, and R is the corresponding multiple correlation coefficient.

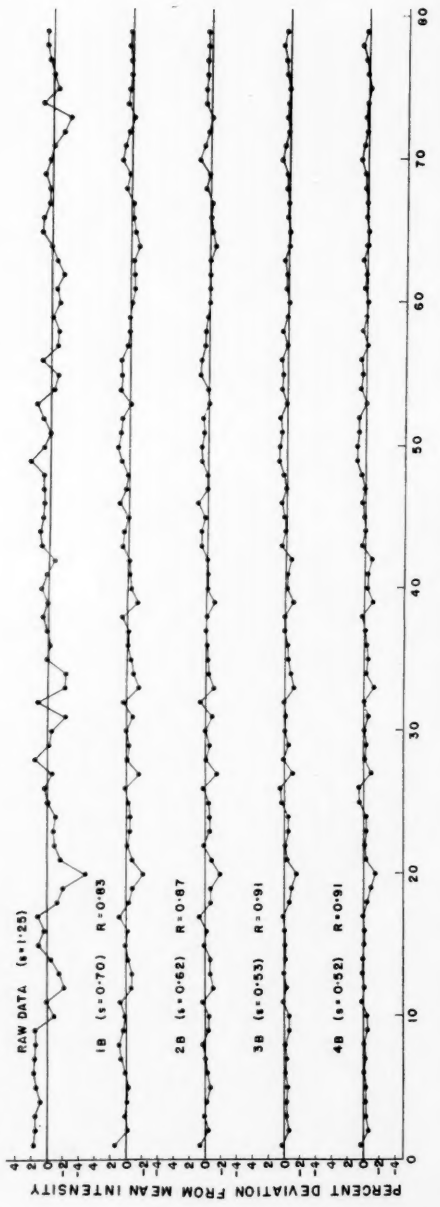


Fig. 2. Results of correcting wide-angle cosmic-ray intensity according to regression coefficients found from preliminary analyses.

From the preliminary calculations it was clear that further five-variate analysis was not warranted because of its extreme additional complexity and doubtful practical value, in agreement with WADA and KUDO (1956).

Table 2. Computational subgroups and time periods associated with final calculations

Equipment period	Telescope*	Time bracket†	Subgroup symbol	Calendar period included
I	A	1	I–A1	1 September 1953–4 April 1954
		2	I–A2	
	B	1	I–B1	
		2	I–B2	
II	A	1	II–A1	5 April 1954–31 October 1955
		2	II–A2	
	B	1	II–B1	
		2	II–B2	
III	A	1	III–A1	1 November 1955–16 February 1956
		2	III–A2	
	B	1	III–B1	
		2	III–B2	
Total	A	1	Total–A1	1 September 1953–16 February 1956
		2	Total–A2	
	B	1	Total–B1	
		2	Total–B2	

* Telescope A—narrow angle; telescope B—wide angle.
† Time bracket 1—0300 UT radiosonde ascent; time bracket 2—1500 UT radiosonde ascent.

Further refinement of the statistical results was desired in order that wide- and narrow-angle regression coefficients might be compared. Because of the general disagreement upon which atmospheric parameters should be used, it was decided to try four "models" involving the seven variables listed in Section 3 and to attempt to determine which model gave the most consistent results for six different data-analysis groups. These groups were determined by the radiosonde flight and periods of instrumental similarity. The models are as follows:

Model	*Included variables*
50 mb DUPERIER (1951)	I_A; I_B; B, H_5, T_{52}
100 mb DUPERIER (1949)	I_A; I_B; B, H_1, T_{12}
200 mb DUPERIER (1956)	I_A; I_B; B, H_2, T_{52}
OLBERT (1953)	I_A; I_B; B, H_1, T_t

Definitions of the variables are found in Section 3.

VOL. 14 1959

In order to investigate possible diurnal differences, the data were grouped by radiosonde flights and all calculations were carried out separately for the 0300 UT and 1500 UT flights. To use the data economically, no attempt was made to select for analysis only those days for which all necessary data were present for both flight periods. For the most part, however, such a grouping of data does represent a time separation of 12 hr, and any non-statistical differences in the results can be considered as possible diurnal effects.

The data-selection criteria were altered slightly from those used in the preliminary calculations in order to take full advantage of the cosmic-ray intensity data. Many radiosonde flights per month reached 100 mb but a considerably smaller number reached 50 mb. The data groups were essentially parallel, but there were, e.g., more 100 mb entries than there were for 50 mb. Likewise, the number of datum periods for the two daily radiosonde flights were only approximately equal for each data group.

The various data subgroups and associated calendar periods are listed in Table 2. In the course of the experiment, some changes in equipment sensitivity were required, and the data were grouped into three calendar periods during each of which only relatively minor changes were made in the equipment. Division into data groups for which the equipment was essentially constant appeared to be an adequate choice for the study of "reproducibility" of regression coefficients. Sixteen computational subgroups were thus established, and a total of sixteen triple-regression calculations were completed for each of the four models. In each case the regression equation was of the form given in equation (1) and the results are tabulated for each model in Tables 3, 4, 5 and 6.

Table 3. Final regression coefficients using 50 mb Duperier model (DUPERIER, 1951). Coefficients include standard deviations. See Table 2 for subgroup identification and refer to text (Section 4) for variables included in calculations.

Subgroup	No. of datum periods	Simple barometric b_0 (%/cm Hg)	Partial barometric b_1 (%/cm Hg)	Partial height b_2 (%/km)	Partial temperature b_3 (%/°C)	R
Total–A 1	438	−0·69 ± 0·19	−0·96 ± 0·17	−3·54 ± 0·42	+0·100 ± 0·031	0·47
Total–A 2	543	−0·85 ± 0·12	−1·13 ± 0·14	−3·95 ± 0·36	+0·156 ± 0·027	0·56
I–A 1	115	−1·30 ± 0·22	−1·42 ± 0·16	−3·35 ± 0·67	+0·222 ± 0·031	0·78
I–A 2	135	−1·48 ± 0·21	−1·48 ± 0·15	−4·53 ± 0·64	+0·189 ± 0·031	0·78
II–A 1	273	−0·65 ± 0·29	−1·06 ± 0·27	−3·30 ± 0·52	+0·084 ± 0·047	0·43
II–A 2	344	−0·50 ± 0·24	−0·79 ± 0·21	−3·86 ± 0·43	+0·158 ± 0·039	0·54
III–A 1	50	−0·49 ± 0·30	−0·48 ± 0·30	−2·60 ± 2·25	+0·104 ± 0·056	0·37
III–A 2	64	−1·50 ± 0·22	−1·43 ± 0·22	−1·80 ± 1·96	+0·129 ± 0·044	0·70
Total–B 1	438	−0·85 ± 0·16	−1·16 ± 0·12	−3·56 ± 0·30	+0·179 ± 0·022	0·67
Total–B 2	543	−0·96 ± 0·12	−1·19 ± 0·09	−3·08 ± 0·22	+0·195 ± 0·016	0·73
I–B 1	115	−1·47 ± 0·20	−1·58 ± 0·05	−2·53 ± 0·21	+0·203 ± 0·009	0·97
I–B 2	135	−1·44 ± 0·14	−1·42 ± 0·08	−3·03 ± 0·34	+0·166 ± 0·017	0·90
II–B 1	273	−0·81 ± 0·21	−1·21 ± 0·19	−3·08 ± 0·36	+0·138 ± 0·023	0·57
II–B 2	344	−1·21 ± 0·16	−1·41 ± 0·11	−1·67 ± 0·23	+0·197 ± 0·021	0·70
III–B 1	50	−1·22 ± 0·21	−1·21 ± 0·16	−3·29 ± 1·23	+0·155 ± 0·031	0·81
III–B 2	64	−1·35 ± 0·12	−1·24 ± 0·10	−4·05 ± 0·90	+0·109 ± 0·020	0·89

Table 4. Final regression coefficients using 100 mb model (DUPERIER, 1949)

Subgroup	No. of datum periods	Simple barometric b_0 (%/cm Hg)	Partial barometric b_1 (%/cm Hg)	Partial height b_2 (%/km)	Partial temperature b_3 (%/°C)	R
Total–A 1	583	-0.74 ± 0.15	-0.91 ± 0.14	-3.03 ± 0.38	$+0.076 \pm 0.022$	0·48
Total–A 2	618	-1.43 ± 0.19	-1.14 ± 0.13	-3.19 ± 0.38	$+0.136 \pm 0.021$	0·56
I–A 1	150	-1.28 ± 0.19	-1.17 ± 0.14	-3.92 ± 0.63	$+0.092 \pm 0.025$	0·74
I–A 2	157	-1.48 ± 0.19	-1.44 ± 0.13	-4.57 ± 0.56	$+0.129 \pm 0.023$	0·81
II–A 1	368	-0.55 ± 0.23	-0.98 ± 0.23	-2.80 ± 0.50	$+0.077 \pm 0.031$	0·44
II–A 2	389	-0.42 ± 0.23	-0.81 ± 0.20	-2.80 ± 0.49	$+0.154 \pm 0.030$	0·52
III–A 1	65	-0.74 ± 0.29	-0.70 ± 0.28	-3.13 ± 1.91	$+0.053 \pm 0.039$	0·42
III–A 2	72	-1.46 ± 0.20	-1.44 ± 0.19	-2.00 ± 1.60	$+0.084 \pm 0.032$	0·51
Total–B 1	583	-0.88 ± 0.13	-1.25 ± 0.10	-3.61 ± 0.27	$+0.076 \pm 0.016$	0·67
Total–B 2	618	-1.57 ± 0.16	-1.24 ± 0.10	-3.69 ± 0.29	$+0.091 \pm 0.017$	0·65
I–B 1	150	-1.52 ± 0.16	-1.43 ± 0.07	-2.75 ± 0.30	$+0.087 \pm 0.012$	0·91
I–B 2	157	-1.46 ± 0.13	-1.41 ± 0.07	-3.25 ± 0.31	$+0.109 \pm 0.013$	0·91
II–B 1	368	-0.99 ± 0.17	-1.37 ± 0.15	-2.36 ± 0.35	$+0.075 \pm 0.022$	0·56
II–B 2	389	-1.10 ± 0.18	-1.37 ± 0.16	-1.94 ± 0.39	$+0.104 \pm 0.024$	0·53
III–B 1	65	-1.16 ± 0.18	-1.12 ± 0.15	-3.96 ± 1.01	$+0.067 \pm 0.021$	0·78
III–B 2	72	-1.33 ± 0.12	-1.28 ± 0.08	-4.89 ± 0.70	$+0.045 \pm 0.014$	0·83

VOL. 14 1959

Table 5. Final regression coefficients using 200 mb model (DUPERIER, 1956)

Subgroup	No. of datum periods	Simple barometric b_0 (%/cm Hg)	Partial barometric b_1 (%/cm Hg)	Partial height b_2 (%/km)	Partial temperature b_3 (%/°C)	R
Total–A 1	438	-0.69 ± 0.19	-0.98 ± 0.17	-3.34 ± 0.41	-0.045 ± 0.041	0·47
Total–A 2	543	-0.85 ± 0.12	-1.12 ± 0.14	-3.70 ± 0.35	$+0.013 \pm 0.034$	0·55
I–A 1	115	-1.30 ± 0.22	-1.34 ± 0.17	-2.88 ± 0.70	$+0.113 \pm 0.048$	0·76
I–A 2	135	-1.48 ± 0.21	-1.49 ± 0.15	-4.69 ± 0.56	$+0.012 \pm 0.040$	0·80
II–A 1	273	-0.65 ± 0.29	-1.02 ± 0.27	-3.06 ± 0.52	-0.060 ± 0.059	0·41
II–A 2	344	-0.50 ± 0.24	-0.89 ± 0.20	-3.73 ± 0.30	$+0.268 \pm 0.034$	0·63
III–A 1	50	-0.49 ± 0.30	-0.49 ± 0.29	-4.31 ± 2.24	-0.066 ± 0.102	0·42
III–A 2	64	-1.50 ± 0.22	-1.47 ± 0.22	$+0.02 \pm 2.68$	$+0.115 \pm 0.102$	0·70
Total–B 1	438	-0.85 ± 0.16	-1.19 ± 0.12	-3.51 ± 0.29	$+0.023 \pm 0.028$	0·68
Total–B 2	543	-0.96 ± 0.12	-1.20 ± 0.08	-3.08 ± 0.20	-0.071 ± 0.020	0·74
I–B 1	115	-1.47 ± 0.20	-1.53 ± 0.07	-1.88 ± 0.27	$+0.135 \pm 0.019$	0·95
I–B 2	135	-1.44 ± 0.14	-1.42 ± 0.07	-3.07 ± 0.29	$+0.050 \pm 0.021$	0·91
II–B 1	273	-0.81 ± 0.21	-1.18 ± 0.19	-2.50 ± 0.36	$+0.020 \pm 0.042$	0·53
II–B 2	344	-1.21 ± 0.16	-1.50 ± 0.09	-2.34 ± 0.14	$+0.246 \pm 0.016$	0·82
III–B 1	50	-1.22 ± 0.21	-1.22 ± 0.17	-2.93 ± 1.28	$+0.038 \pm 0.059$	0·80
III–B 2	64	-1.35 ± 0.12	-1.28 ± 0.07	-8.85 ± 0.84	-0.226 ± 0.032	0·95

Table 6. Final regression coefficients using arithmetic mean temperature of troposphere (OLBERT, 1953)

Subgroup	No. of datum periods	Simple barometric b_0 (%/cm Hg)	Partial barometric b_1 (%/cm Hg)	Partial height b_2 (%/km)	Partial temperature b_3 (%/°C)	R
Total–A 1	583	−0·74 ± 0·15	−1·24 ± 0·19	−0·42 ± 1·27	−0·117 ± 0·043	0·47
Total–A 2	618	−1·43 ± 0·19	−1·90 ± 0·18	+2·83 ± 1·27	−0·258 ± 0·043	0·56
I–A 1	150	−1·28 ± 0·19	−1·86 ± 0·20	+2·51 ± 1·74	−0·244 ± 0·054	0·76
I–A 2	157	−1·48 ± 0·19	−2·32 ± 0·20	+4·07 ± 1·97	−0·322 ± 0·061	0·81
II–A 1	368	−0·55 ± 0·23	−1·10 ± 0·27	−1·95 ± 1·51	−0·056 ± 0·055	0·42
II–A 2	389	−0·42 ± 0·23	−1·65 ± 0·27	+3·13 ± 1·68	−0·265 ± 0·059	0·51
III–A 1	65	−0·74 ± 0·29	−1·63 ± 0·57	+8·36 ± 6·80	−0·295 ± 0·159	0·44
III–A 2	72	−1·46 ± 0·20	−1·75 ± 0·26	+0·44 ± 2·40	−0·117 ± 0·059	0·70
Total–B 1	583	−0·88 ± 0·13	−1·63 ± 0·37	−0·48 ± 2·52	−0·135 ± 0·086	0·67
Total–B 2	618	−1·57 ± 0·16	−1·84 ± 0·13	+1·20 ± 0·98	−0·203 ± 0·033	0·66
I–B 1	150	−1·52 ± 0·16	−2·07 ± 0·22	+3·26 ± 1·90	−0·228 ± 0·058	0·45
I–B 2	157	−1·46 ± 0·13	−2·31 ± 0·10	+5·98 ± 0·97	−0·334 ± 0·030	0·93
II–B 1	368	−0·99 ± 0·17	−1·59 ± 0·19	−0·59 ± 1·13	−0·087 ± 0·039	0·55
II–B 2	389	−1·10 ± 0·18	−2·07 ± 0·22	+3·32 ± 1·33	−0·224 ± 0·046	0·54
III–B 1	65	−1·16 ± 0·18	−1·88 ± 0·31	+5·16 ± 3·70	−0·241 ± 0·087	0·77
III–B 2	72	−1·33 ± 0·12	−1·48 ± 0·11	−3·33 ± 1·06	−0·068 ± 0·026	0·91

In Tables 7 and 8 are presented the standard deviations of the atmospheric parameters appropriate to the various calculations. The regression coefficients listed in Table 9 are the weighted means of the subgroup coefficients (Tables 3–6) and are separated only with regard to counter-telescope geometry. It is important to note that the standard deviations quoted in Table 9 are the standard deviations of the aggregate of the six subgroup regression coefficients found for each term of the regression equation. The latter were used as final criteria of experimental reproducibility of the subgroup coefficients, since if the regression coefficients

Table 7. Standard deviations of variables for 50 mb and 200 mb Duperier analyses

Group and flight	Datum periods	I_A*	I_B†	B (cm Hg)	H_5 (km)	H_2 (km)	T_{52} (°C)
Total–1	438	15·5	38·2	0·531	0·227	0·306	3·0
Total–2	543	16·3	35·2	0·565	0·229	0·299	3·0
I–1	115	12·5	29·3	0·636	0·159	0·235	3·4
I–2	135	14·0	31·7	0·677	0·169	0·243	3·4
II–1	273	15·8	35·2	0·449	0·239	0·307	2·6
II–2	344	15·8	29·1	0·463	0·241	0·301	2·6
III–1	50	9·5	24·2	0·584	0·077	0·142	3·1
III–2	64	11·9	25·1	0·683	0·080	0·133	3·5

* Register counts/hr. Multiply by 4 to get total.
† Register counts/hr. Multiply by 16 to get total.

Table 8. Standard deviations of variables for 100 mb Duperier and Olbert analyses

Group and flight	Datum periods	I_A*	I_B†	B (cm Hg)	H_2 (km)	T_{12} (°C)	T_t (°C)
Total–1	583	15·2	37·6	0·541	0·239	4·1	7·5
Total–2	618	16·3	40·8	0·572	0·237	4·1	7·4
I–1	150	12·1	29·1	0·633	0·169	4·2	5·7
I–2	157	13·8	32·1	0·674	0·177	4·3	6·0
II–1	368	15·3	34·5	0·466	0·247	3·9	7·5
II–2	389	16·0	37·7	0·474	0·241	3·8	7·3
III–1	65	10·7	24·2	0·588	0·089	4·4	4·2
III–2	72	11·8	25·9	0·703	0·086	4·3	3·9

* Register counts/hr. Multiply by 4 to get total.
† Register counts/hr. Multiply by 16 to get total.

have definite physical meaning, one would expect that they could be calculated for a particular experimental apparatus (perhaps eventually predicted from the theory), making it unnecessary to recompute them from time to time. In other words, one hopes that they would be constants of the apparatus, and the atmospheric models were tested in this regard (see Section 5 below) by making comparisons of the standard deviations of the subgroup regression coefficients.

Table 9. Weighted means of the partial regression coefficients. The standard deviation of the aggregate of the component six subgroup coefficients is given for each weighted mean coefficient. See Tables 3–6 for component values.

Model	Telescope	Partial barometric b_1 (%/cm Hg)	Partial height b_2 (%/km)	Partial temperature b_3 (%/°C)
50 mb Duperier*	A	−1·01 ± 0·30	−3·77 ± 0·66	+0·131 ± 0·047
100 mb Duperier†	A	−1·03 ± 0·24	−3·11 ± 0·70	+0·107 ± 0·036
200 mb Duperier*	A	−1·06 ± 0·38	−3·54 ± 2·88	−0·013 ± 0·140
Olbert†	A	−1·60 ± 0·39	+1·25 ± 2·84	−0·190 ± 0·104
50 mb Duperier*	B	−1·18 ± 0·13	−3·29 ± 0·75	+0·188 ± 0·030
100 mb Duperier†	B	−1·24 ± 0·07	−3·65 ± 0·79	+0·083 ± 0·018
200 mb Duperier*	B	−1·20 ± 0·15	−3·27 ± 1·61	−0·029 ± 0·130
Olbert†	B	−1·73 ± 0·27	+0·38 ± 2·67	−0·169 ± 0·087

* Total datum periods 981.
† Total datum periods 1201.

5. Accuracy of the Atmospheric Parameters

Microbarograph data were supplied by the United States Weather Bureau, Lincoln, Nebraska. The standard error of measurement of the barometric

pressure was approximately 0·0127 cm Hg. A typical standard deviation of the barometric pressure is about 0·50 cm Hg, and the ratio of the standard error of measurement to the standard deviation is approximately 0·025.

The quality of the radiosonde data is more difficult to analyse. The altitude intervals used in this investigation range from near sea level to the height of the 50 mb level, and no general statement is possible which will apply to all levels. Not a great deal has been published concerning this, but a report of a series of paired radiosonde flight compatibility tests was made by the United States Weather Bureau in 1951 (United States Department of Commerce, 1951). A summary of subsequent work was obtained from Thickstun (1956), who also discussed radiosonde similitude tests made at the factory, where the instruments were placed in a chamber in which the pressure and temperature were varied at a rate closely following that experienced in actual flight. The conclusions, based on both of these tests, are as follows:

(1) The tests showed a marked failure of all types of radiosondes to repeat their own recordings.

(2) There is a wide scatter of differences between radiosonde pairs soon after leaving the ground (ascribed to the inconsistency noted above).

(3) The differences are so large as to mask other comparisons, e.g. external effects, such as those due to solar radiation.

Thickstun estimated quantitatively on the basis of a study on progress at the time (1956) and the 1951 report that, for the most part, errors in the height of the 100 mb level are not associated with the individual instruments but with certain operating difficulties at the various stations. Maximum uncertainties in the height of the 100 mb level are about ± 20 m. The standard error of measurement of the 100 mb level would be approximately 10 m, and in the present computational work isobaric heights were rounded to the nearest 10 m (within ± 5 m of the reported value). The ratio of the standard error of measurement to a typical standard deviation (Table 9) would be about 0·05 for the 100 mb level, but it goes as high as approximately 0·12 for Group III-2.

Comparable data on the 50 mb heights are not available, but somewhat higher uncertainties are to be expected. According to Thickstun, it will be a rare case when the height of the 50 mb surface will be in error by more than 60 m. A typical value of the ratio of the standard error of measurement of height of the 50 mb level to the standard deviation is estimated to be 0·15, but it approaches 0·40 for Group III-1.

The temperature measurements are considerably less reliable. The maximum differences between any pair of radiosondes from twenty experimental soundings was 1·5°C, starting from zero at the surface and increasing to the maximum value at the end of the flight. The absolute value of the temperature may be incorrect by as much as 2°C and occasionally more. The standard error for all twenty soundings in the above test was 0·55°C. It is probable that more than average care was taken in these Weather Bureau experiments, and hence one might expect larger standard errors from routine station soundings. On this basis, the ratio of the standard error of measurement to the standard deviation of the mean temperature between the 100 and 200 mb levels is estimated to be about 0·20 and

that for the mean temperature between the 50 and 200 mb levels to be approximately 0·33. The corresponding ratio for the arithmetic mean temperature of the troposphere would be considerably lower, and it is approximately 0·10 or even less.

6. Comparison Amongst Models Used

Examination of the standard deviations of the partial barometric coefficients b_1 in Table 9 indicates that for telescope A there is little difference amongst models. For telescope B a similar comparison shows the 100 mb Duperier model to be slightly more reliable than the 50 mb Duperier; the standard deviation (as well as the regression coefficient) is somewhat larger for the wide-angle Olbert calculation.

A similar comparison of the standard deviations of the partial height coefficients b_2 for telescope A reveals that those for the 50 mb and 100 mb Duperier models are nearly equal and also that they are approximately one-fourth the size of the standard deviations found for the 200 mb Duperier and Olbert models. For telescope B the 50 mb and 100 mb Duperier models give coefficients which are considerably more reproducible than those of the 200 mb Duperier and Olbert models.

The standard deviations of the temperature coefficient b_3 for telescope A are also smaller for the 50 mb and 100 mb analyses than those for the 200 mb Duperier and Olbert cases. A similar comparison holds for telescope B. Thus we see that the temperature coefficients b_3 are also experimentally more reproducible for the 50 mb and 100 mb Duperier models than for the 200 mb Duperier and Olbert models.

VOL. 14 1959

The foregoing leads one to conclude that the 50 mb and 100 mb Duperier models give more reproducible results, and hence, by our criteria, are superior to the 200 mb Duperier and Olbert analyses. We now wish to compare the 50 mb and 100 mb models to see whether or not a distinct choice can be made between them. From examination of Table 9 and the arguments given above, it is seen that, in terms of the partial barometric coefficients b_1, the 100 mb Duperier model is a slight favourite, but both the 50 mb and the 100 mb models are nearly equally reproducible in terms of the height coefficients b_2 and the temperature coefficients b_3. The 50 mb Duperier model is somewhat superior when the standard deviations of the temperature coefficients are taken relative to the coefficients themselves. Thus, if these were the only considerations, one could not decide between the 50 mb and the 100 mb Duperier models. On the basis of the discussion in Section 5 of the accuracy of the atmospheric parameters, however, one would expect that if the two models were equally good physically, then the standard deviation for the 50 mb Duperier height coefficients would be nearly three times those for the 100 mb cases, and the 50 mb temperature coefficients would have nearly twice the standard deviation of the 100 mb values, taking the same number of datum periods for each. Since the actual standard deviations are nearly the same, and a smaller number of datum periods were involved in the 50 mb Duperier analyses, we must conclude that the 50 mb Duperier model is the best choice of the four statistical models for making atmospheric corrections to both narrow- and wide-angle cosmic ray intensity data.

7. Comparison Amongst Different Flights and between Counter Telescopes

We now wish to compare the separate results for the two radiosonde flights to see if any diurnal differences exist between results of the same calculational group. These data were generally separated in time by 12 hr. Only the results of the 50 mb and 100 mb Duperier analyses will be discussed in any detail, although similar results were obtained for the other analyses also. These results are conveniently grouped for comparison purposes in Tables 3 and 4. A comparison of the partial temperature coefficients b_3 for flights 1 and 2 indicate that in nearly every case the results are statistically different or barely overlap in the extreme limits of their standard deviations. A similar situation exists for the partial height coefficients b_2 and the partial barometric coefficients b_1, which are in some cases radically different between the two flight periods. From these semiquantitative arguments the results for the two radiosonde flights appear to be different.

In addition to this semiquantitative indication of diurnal differences, an F-test (Rao, 1952) was used to test the hypothesis that there was equality between corresponding regression coefficients obtained from the analysis of each of the radiosonde flights. It was possible in every one of the groups tested to reject the hypothesis that they were equal at the 1 per cent level of significance. This result must be interpreted with caution, however, since a similar result could occur if the coefficients were apparently equal but the data had come from statistically different populations. Such a situation can exist in an investigation of this sort, but the F-test results provide further evidence that the results are statistically different for the two radiosonde flights. Thus we conclude that there is little question as to the statistical reality of the difference, but the crucial question remains as to why such a difference exists.

A first possibility is that the radiosonde errors vary in some diurnal way. Incomplete radiation correction of the upper-altitude temperature measurements would be a possible explanation. Second, it is possible that the linear regression equation (1) used is not correct. Higher order terms may actually be required to account fully for the atmospheric effects. Third, it may arise from a diurnal variation in the cosmic ray primary spectrum, but the evidence is far from conclusive. Further study of this apparent diurnal variation will be undertaken.

A comparison similar to that above can be applied to the simultaneous results for the narrow- and wide-angle telescopes, as presented in Table 9. It is noted that a difference does exist between the results as a whole for the narrow-angle telescope and the wide-angle telescope. Application of the above F-test showed them to be statistically different by that criterion as well.

8. General Comparison with Results of Other Investigators

Not many workers have used a reference level of 50 mb. Dolbear and Elliot (1951) used this level in calculations with intensity data from a pair of inclined counter telescopes, which makes comparison with the present work difficult. The general agreement with other workers (summary in Bachelet and Conforto,

1956) is excellent for the temperature coefficients and good for the barometric coefficients. The average Duperier height coefficient is somewhat larger than those found in the present work, but the standard deviations are so large as to make the results reasonably compatible.

A considerably larger number of investigators have used the 100 mb Duperier analysis. In addition to the work of authors previously mentioned, the results of COTTON and CURTIS (1951), DAWTON and ELLIOT (1953), TRUMPY and TREFALL (1953) and TRUMPY (1954), may be compared. There is no clear agreement as to the temperature coefficients, but the general agreement of the present work with the results of DUPERIER is excellent.

Not many results are available for the 200 mb Duperier analysis, but, again, the present work is seen to agree with that of DUPERIER.

A somewhat similar situation exists for the Olbert analysis. The work of WADA and KUDO (1954) is in good agreement with the present results except for the temperature coefficient.

In all of the work cited it is relevant to note the approximate ratio of the standard error (or, in the present work, the standard deviation) to the magnitude of the corresponding regression coefficient. These ratios vary from about 0·13 to more than 1·00 for the temperature coefficient, with about 0·35 being a typical value. The ratios for the height coefficient vary from 0·12 to more than 1·00, with about 0·15 being typical. Such fluctuations would be expected in view of the uncertainties of measurement of atmospheric parameters.

VOL. 14 1959

9. COMPARISON WITH THEORY

When the 50 mb Duperier results in Table 9 are compared with the predictions of TREFALL (1955b), there is apparent disagreement. For the absorber used in the present work, TREFALL's predicted value of the 50 mb temperature coefficient (0·215%/°C) is somewhat higher than the experimental values. TREFALL showed, however, that the temperature coefficient for mesons arriving from a zenith direction θ is given by the approximate formula

$$b_3(\theta) = b_3(\mathrm{O}) + k(\ln \cos \theta) \qquad (2)$$

where k is a slowly-varying function of the cut-off momentum of the recorder and has a value of approximately 0·12%/°C for the results to be considered here. The term $b_3(\mathrm{O})$ is the coefficient for vertically incident particles. The theoretical curve shown by TREFALL is for vertically-incident particles, and equation (2) gives the converted value for the wide-angle telescope of approximately 0·22%/°C, in good agreement with the Trefall theory. The present narrow-angle result seems low, but the standard error is such that it is not inconsistent with the theory.

One notes good agreement upon comparing the results in Table 9 with similar theoretical predictions for the temperature coefficient in the case where the 100 mb reference level is taken. The result for the wide-angle telescope is about 0·024%/°C lower than that for the narrow-angle telescope, in excellent agreement with equation (2).

Further, it is interesting to compare the ratio of the 50 mb temperature coefficient with the 100 mb temperature coefficient according to the TREFALL

theory and the experimental work of Duperier. Trefall predicts a ratio of 2·5, while Duperier (1951) obtained 1·84 for a narrow-angle telescope and 1·91 for one with a wider aperture. The present work gives 1·23 for the narrow-angle results and 2·26 for the wide-angle results. Considering the standard errors involved this is not incompatible with the theory, and it is in excellent agreement with the experimental work of Duperier.

No theoretical work is available which makes possible a similar comparison of 200 mb experimental results. As noted earlier, the 200 mb scheme is definitely inferior to the 50 mb and 100 mb Duperier models, and it will not be discussed further here.

The results in Table 9 do not agree with all of those predicted by Olbert (1953), who calculated the atmospheric effects on the basis of vertically-incident mesons. While comparison is not quite feasible for the wide-angle telescope, the results for the narrow-angle telescope should afford a test of the vertical case analysed by Olbert. The results in Table 9 show that although the barometric coefficients agree well, both of the temperature coefficients found on the present work are about three times too large, and the height coefficients are of the wrong sign when compared with Olbert's predictions. This must be interpreted cautiously, however, since the standard deviations are very large relative to the magnitude of the regression coefficients. The experimental results of Wada and Kudo (1954) are similar to these. Thus it is seen that the Olbert theory does not predict the measured values for the height and temperature cofficients. In addition the calculation based on the Olbert model does not give the experimental reproducibility of the 50 mb and 100 mb Duperier models, leading one to reject it on those experimental criteria as well. It is possible that use of a weighted mean tropospheric temperature would yield results in better agreement with Olbert.

It is noted from Table 9 that the height coefficients do not seem to depend upon the aperture of the recorder. From consideration of the three Duperier-type analyses, it appears that the height coefficients b_2 are not statistically different. As Wada and Kudo (1954) point out, there are two competing processes affecting the intensity of obliquely incident particles; i.e. the decay effect increases because of increasing path length to the recorder, but the effective mean momentum of the μ-mesons traversing obliquely through the atmosphere is higher, and is inversely proportional to the decay probability. Since the magnitudes of the height coefficients b_2 remain essentially constant, the gain effect due to increased momentum seems to cancel the loss due to the increased decay effect resulting from longer average paths. No theoretical work is available for a quantitative estimate of the individual effects. Wada and Kudo have clearly indicated the extreme complexity of the omnidirectional problem and the ultimate necessity of its analysis in terms of the complicated meson-production spectrum together with its altitude dependence.

Further reference to Table 9 reveals that the partial barometric coefficients b_1 have an apparent zenith-angle dependence. It is interesting to note that the barometric coefficients do not depend upon the reference level for the three Duperier-type analyses but do depend upon the aperture of the recorder. The difference between narrow- and wide-angle telescopes is in each case greater than

one would expect from increased path length alone, possibly indicating that the magnitude of the μ-e decay effect increases with zenith angle also. No quantitative theoretical work has been done on the zenith-angle dependence of the barometric effect. The work of TREFALL (1955c) considers only vertically-travelling μ mesons, and hence can only approximately predict the observed experimental values. His predicted value of the partial barometric coefficient for the cut-off momentum of the equipment used in the present work is approximately $-1{\cdot}60\%$/cm Hg. This is considerably larger than the results of the narrow-angle barometric coefficients in Table 9, which are about $-1{\cdot}03\%$/cm Hg for all three narrow-angle Duperier analyses. The wide-angle results are low also, being approximately $-1{\cdot}24\%$/cm Hg.

The contribution of the absorption effect to the partial barometric coefficient as derived by TREFALL, is not directly a function of station elevation, but depends upon the change in station barometric pressure, the cut-off momentum of the recorder and the differential and integral momentum spectra. The latter do depend upon station elevation, but the ratio of the two appears in TREFALL's absorption expression, and the net change with the station elevation would be slight between sea level and 350 m altitude.

The contribution due to the decay of μ-mesons is more complicated, but to a first approximation, it would decrease with the increasing elevation since the mean momentum of the μ-mesons is a strong factor in the denominator of TREFALL's expression. This contribution, however, is not sufficient for the Lincoln, Nebraska, elevation of 350 m to explain the 50 per cent increase required to bring the results of the present work into agreement with the value predicted by TREFALL. The standard deviations are such that the above results are not incompatible with the theory, but the agreement cannot be considered good. Again it is interesting to note the agreement with the present work and that of DUPERIER both for zenith-angle dependence and magnitude of the partial barometric coefficient b_1.

10. GENERAL CONCLUSIONS

First, the extent to which the cosmic ray meteorological effects can be eliminated at the present time is limited by the accuracy of the radiosonde data available. GLOKOVA (1954) has reported similarly.

Second, the 50 mb Duperier-type analysis yields results which are, in principle, more reproducible than those of the 100 mb analysis, and the former are equally reproducible in practice despite increasing radiosonde errors at the higher levels. Both analyses are superior to the 200 mb Duperier and Olbert type analyses.

Third, separate analysis of the data for the 0300 UT and the 1500 UT radiosonde flights gave statistically different results indicating a possible diurnal variation of the atmospheric coefficients.

Fourth, the results for each type of analysis are in general agreement with the experimental work of other investigators. Agreement with the work of Duperier is especially good.

Fifth, the present results are in good agreement with the theoretical partial temperature-effect predictions of TREFALL and are in poor agreement with his

VOL. 14 1959

predicted partial barometric effects; they generally disagree with the work of OLBERT when a simple arithmetic-mean temperature of the troposphere is used.

Acknowledgements—We are indebted to the University of Nebraska Research Council for providing equipment funds. One of us (W. R. F.) is grateful for three grants from the Research Corporation, and to Nebraska Wesleyan University for a leave of absence to do this work. Messrs. B. ANSPAUGH, W. LANG, J. PARK and R. CYR did much of the computational work, and Dr. B. FRENCH gave helpful advice on problems of numerical analysis. The generous aid of the personnel of Lincoln, Nebraska, Station of the United States Weather Bureau is also acknowledged.

REFERENCES

BACHELET F. and CONFORTO A. M.	1956	*Nuovo Cim.* **4,** 1479.
CHASSON R. L.	1954	*Phys. Rev.* **96,** 1116.
CHASSON R. L. and FRENCH W. R.	1955	Proceedings of the International Congress on Cosmic Rays, Guanajuato. Unpublished.
COTTON E. S. and CURTIS H. O.	1951	*Phys. Rev.* **84,** 840.
DAWTON D. I. and ELLIOT H.	1953	*J. Atmosph. Terr. Phys.* **3,** 295.
DOLBEAR D. W. N. and ELLIOT H.	1951	*J. Atmosph. Terr. Phys.* **1,** 215.
DORMAN L. I. *et al.*	1953	*Zh. Eksp. Teor. Fiz.* **26,** 537.
DORMAN L. I. and FEINBERG E. L.	1955	Proceedings of the International Congress on Cosmic Rays, Guanajuato. Unpublished.
DUPERIER A.	1949	*Proc. Phys. Soc. Lond.* A **62,** 684.
DUPERIER A.	1951	*J. Atmosph. Terr. Phys.* **1,** 296.
DUPERIER A.	1956	Private communication.
GLOKOVA E. S.	1954	Report of Moscow Conference on Cosmic Ray Physics. Unpublished.
MCNISH A. G.	1936	*Terr. Magn. Atmos. Elect.* **41,** 37.
OLBERT S.	1953	*Phys. Rev.* **92,** 454.
RAO C. R.	1952	*Advanced Statistical Methods in Biometric Research.* Wiley, New York.
THICKSTUN W. R.	1956	Private communication.
TREFALL H.	1955a	*Proc. Phys. Soc. Lond.* A **68,** 625.
TREFALL H.	1955b	*Proc. Phys. Soc. Lond.* A **68,** 893.
TREFALL H.	1955c	*Proc. Phys. Soc. Lond.* A **68,** 953.
TREFALL H.	1957	*Physica* **23,** 65.
TRUMPY B. and TREFALL H.	1953	*Physica* **19,** 636.
TRUMPY B.	1954	*Univ. Bergen Årb. Naturv. R.* No. 3.
UNITED STATES DEPARTMENT OF COMMERCE	1951	*A Study of the Report on Radiosonde Compatibility.* Tests made at Oklahoma City, U.S. Weather Bureau, June, 1951.
WADA M. and KUDO S.	1954	*J. Sci. Res. Inst., Tokyo* **48,** 71
WADA M. and KUDO S.	1956	*J. Sci. Res. Inst., Tokyo* **50,** 1.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 19 to 30. Pergamon Press Ltd. Printed in Northern Ireland

The limitations of narrow-beam radio equipments in the detection of weak meteor showers

A. A. Weiss
Division of Radiophysics, C.S.I.R.O., at Department of Physics,
University of Adelaide, Australia

(*Received* 14 *May* 1958)

Abstract—At the echo rates practicable for a routine survey of meteor activity using a narrow-beam radio equipment, it is shown that the chief limitation to the recognition of weak shower activity is the fluctuations in the rate of detection of background (sporadic) meteors. Examination of echo rates obtained with the Adelaide 67 Mc/s equipment for December 1956 and January 1957 confirms that over short intervals of time the background fluctuations are random. A significance test, intended for use as a search method for weak shower activity and non-random fluctuations in the background activity, is then developed. This test, based on echo rates, is applied to the data for December 1956 and January 1957.

1. Introduction

The two principal requirements of a radio equipment specifically intended for routine surveys of meteor shower activity are an adequate echo rate and good resolution as between showers active simultaneously. These objectives are met by the use of high-gain narrow-beam aerial systems (Aspinall *et al.*, 1951).

Two approaches have been made to the problem of the measurement of shower radiant co-ordinates with this type of equipment. The first, due to Clegg (1948) and to Aspinall *et al.* (1951), is suitable for low echo rates and requires the plotting of the slant range and time of occurrence of individual echoes, followed by location of groupings of echoes characteristic of showers (see Fig. 1). At high echo rates the time consumed in such individual echo plots may become prohibitive, and Keay (1957) has shown how accurate radiant co-ordinates may be derived from partial rate plots, i.e. from the rate of echoes received from within narrow limits of slant range.

None of these authors has considered in any detail the problem of recognition, as opposed to measurement of radiant co-ordinates, of weak showers in the presence of the background (sporadic) meteor activity. Most shower radiants lie close to the ecliptic and at elongations of less than 120° from the apex of the earth's way. As the great majority of sporadic meteor radiants are also located in this region the detection of a meteor shower at a time of low background activity is uncommon. In searching for weak showers it must always be borne in mind that the majority of the echoes are from the background; the groupings characteristic of showers do not stand out clearly, and the subjective element is large.

It is at once apparent that the discrimination, for showers as against sporadic meteors, of the type of radio equipment under discussion is inferior to visual and photographic techniques. The latter has a radio counterpart in the multi-station equipment of Gill and Davies (1956), which measures orbits of individual meteors. However, the simpler radio equipment is more useful for routine surveys. It is therefore important to establish the limits of the discrimination of these equipments for meteor showers, and to devise tests for shower activity from which the

subjective element has been eliminated, as far as this is possible. These are the objects of the present paper. As a corollary, proposals for handling data collected in a survey of meteor activity at Adelaide during 1957, with an echo rate of from 500 to 1000 per day, are outlined.

It will be shown that fundamental limitations to the recognition of weak shower activity in records obtained with narrow-beam radio equipments are imposed by random fluctuations in the background activity and by the looseness of some of the weaker streams. The requirement of objectivity can be met, in some degree, by a statistical search method based on total rates; but a constant threshold shower rate, below which recognition is uncertain, cannot then be maintained in the presence of the strong diurnal variation in the background activity (ratio exceeding 20 : 1 in the present survey).

2. The Distinction Between Shower and Background Activity

In this paper the term "meteor shower" is restricted in application to certain periods of increased meteor activity which recur on definite dates. These increases may be large or small, and may persist for intervals extending from a few hours (Giacobinids) to several weeks (Perseids). The meteors of each shower proceed from a definite radiant area, which may be compact or diffuse, and constitute a homogeneous velocity group. Thus the concept of a meteor shower is here associated with a stream of particles moving in a definite orbit round the sun. The numerous published lists of meteor showers all agree on the more intense showers, which may be classed as the major showers. Unfortunately, although perhaps not unexpectedly, there is little agreement as regards the weaker, minor, showers.

After allowing for the radiants of shower meteors, which may be regarded as singularities, the general background distribution of meteor radiants is far from isotropic. Three diffuse concentrations, centred on the sun, the anti-sun and the apex of the earth's way, have been recognized (Hawkins, 1956; Weiss, 1957). These concentrations, which persist throughout the year, are apparent rather than real, in that they are formed by compounding the motion of the earth with the motion of meteor particles whose heliocentric radiant distribution is largely independent of solar longitude. The concept of streams of particles moving in definite orbits is not involved, and these concentrations cannot be considered as showers in the sense of the previous paragraph.

The distinction here drawn between showers and concentrations in the background radiant distribution has a clear physical significance. Provided the shower radiant is not too diffuse, the responses of narrow-beam radio equipments to the two types of meteor distribution are quite different. Should the radiant become very diffuse, the distinction between shower and background can no longer be maintained unless the shower activity is high. A weak shower with a very diffuse radiant will necessarily be allotted to the background activity.

If a meteor shower with a compact radiant area is observed with a narrow-beam radio equipment, and the range and time of detection of each echo is plotted, it will be found that all the shower echoes fall within a range-time envelope whose shape depends on the aerial beam-width and orientation, the latitude of the observing station and the declination of the radiant. Typical range-time envelopes

VOL
14
195

for each of the Adelaide aerials, drawn after the manner described by CLEGG (1948), are shown in Fig. 1; details of the aerial system will be found in WEISS (1955).

Showers will usually be observed against a considerable background activity, and these groupings of echoes characteristic of showers must be accompanied by an increase in the echo rate. Except for radiants close to the south celestial pole these rate increases persist for less than 2 hr. If it is accepted that in a high-rate survey range-time plots of all echoes are impracticable, than a practical method of searching for weak showers must be based upon echo rates alone.

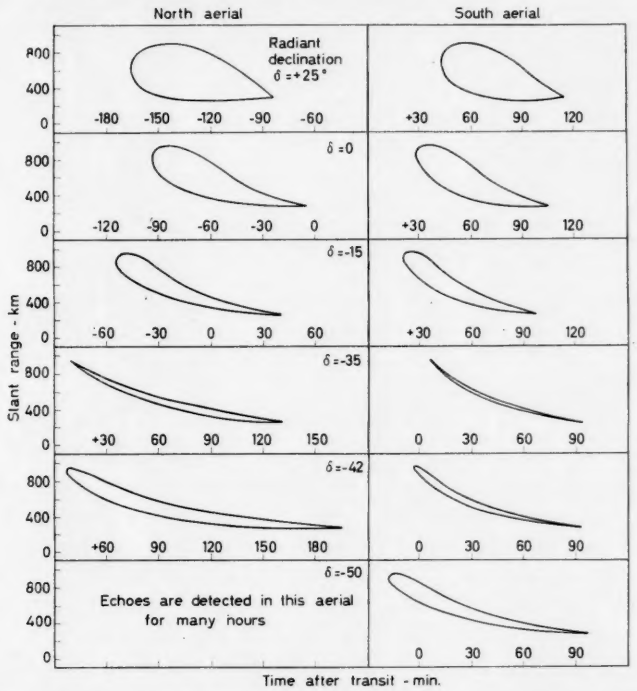


Fig. 1. Range-time envelopes for the Adelaide narrow-beam equipment. North aerial directed to azimuth 35½° S of E, South aerial to 14° N of E.

Now it is well known that the background meteor rate is subject to diurnal and seasonal variations, and also to short-lived statistical fluctuations (Section 4), and mere inspection of echo rates is quite inadequate for the recognition of weak showers. The question then arises, what is to be taken as an increase in the echo rate sufficiently significant for classification as a period of suspected shower activity?

Before considering this question, let us pause to examine the periodic structure of partial rate plots based on short time-intervals.

3. The Random Content of Partial Rate Plots

A typical partial rate plot is illustrated by the lower full line of Fig. 2. This plot gives the number of echoes, per 10 min interval, detected within the limits of

slant range 400–600 km; raw counts, illustrated by the dashed line, are smoothed in a running group of three with the centre interval counted twice. Total counts per $\frac{1}{2}$ hr interval are also shown. With the exception of a few meteors at ranges greater than 900 km, the total count includes all meteors detected within the major lobe of the aerial; no smoothing has been applied. The high rates near midnight are due to the Geminid shower.

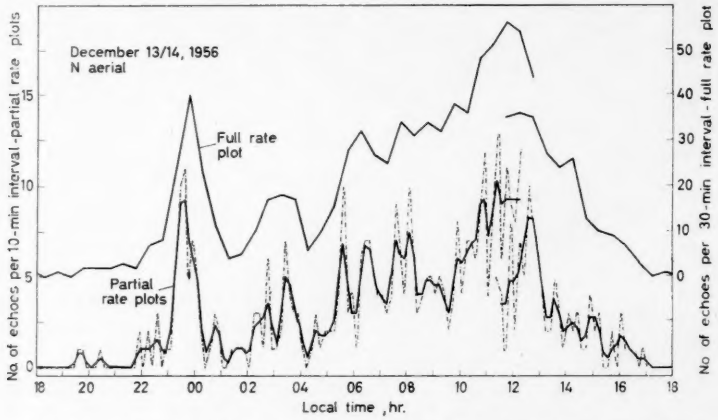


Fig. 2. Typical partial rate and full rate plots of meteor echoes. The partial rate plot includes only echoes detected between 400–600 km slant range; ——— smoothed; -·-·-·-·- unsmoothed. Full rate plot is unsmoothed.

From a casual inspection, using the Geminids as a guide, it would appear possible that some of the other peaks in the partial rate plot could represent shower activity. However, the shape and duration of the peaks is strongly influenced by the smoothing process, and the unsmoothed plot is not so suggestive of shower activity. Whilst some measure of smoothing is both necessary and justifiable if these plots are to be used for radiant *measurement*, it may prove very misleading in radiant *identification*. This is a situation commonly encountered in the interpretation of geophysical data.

Despite the undesirability of smoothing, it is rewarding to follow up the supposition that at least the stronger peaks in the smoothed partial rate plots are to be associated with shower activity. To place the argument on a quantitative basis, there is given in Fig. 3(a) a frequency distribution showing the relative frequency of occurrence of the intervals of time separating successive maxima, and successive minima, in the meteor counts for December 1956. The usual χ^2-tests showed no significant differences as between maxima and minima and as between north and south aerials. It will be seen that the most probable interval separating successive peaks, and successive troughs, is from 40 to 60 min, which is just the interval of time required for the passage of a shower radiant through the aerial beam in the selected range interval of 400–600 km (Fig. 1). This strengthens the inference that many of the peaks in partial rate plots could be due to shower activity, and this inference is made still more plausible by the evident tendency for peaks to repeat at or near the same time from day to day.

But the element of chance coincidence must not be overlooked. For example, if a coincidence in time is scored when the peaks of activity on two successive runs on the same aerial are separated by one interval (10 min) the probability that this coincidence could be due to chance is $\frac{1}{2}$. This is a large probability, but even this method of scoring a coincidence is too stringent for the Geminids, which on one occasion exhibited a displacement of 20 min in time of peak activity.

Moreover, frequency distributions very similar in form to Fig. 3(a) for meteors can be generated by random numbers smoothed in the same manner as the meteor counts. Two examples appear in Figs. 3(b) and 3(c). Fig. 3(b) was generated by allotting equal probability of selection to each random digit from 0 to 9, and

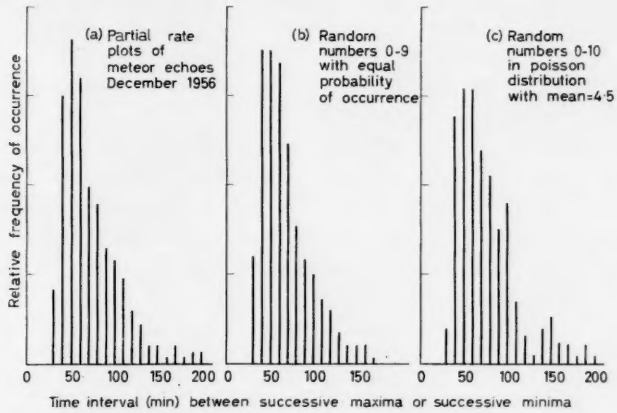

(a) Partial rate plots of meteor echoes December 1956
(b) Random numbers 0-9 with equal probability of occurrence
(c) Random numbers 0-10 in poisson distribution with mean=4·5
Relative frequency of occurrence
0
50
100
150
200
Time interval (min) between successive maxima or successive minima

Fig. 3. Frequency distributions generated by partial rate plots of meteor echoes, and by two sets of random numbers.

Fig. 3(c) by allotting to each random digit from 0 to 10 the probability of selection appropriate to the Poisson distribution with mean $m = 4{\cdot}5$. Some allowance must be made for the diurnal variation in the meteor echo rate, which tends to cause long separations between successive maxima near 1800 hours LT when the meteor rate is low, but this will not destroy the general similarity between the distributions for meteors and for random numbers. This similarity, whose statistical significance is indicated by χ^2 tests, suggests that there is, in fact, a large random component in the meteor counts, and this is confirmed by more detailed comparisons which have been made. These will not be enumerated here. It is sufficient to remark that if the random numbers are divided into groups of 144, to simulate a day's meteor observations, and examined for repetition of features at the same "time" from "day" to "day", then "showers" can be generated with equal facility by random numbers and by actual meteor counts.

The conclusion must be that the great majority of the peaks in the partial rate plots are due to chance fluctuations in the meteor rate. Even taking into account the amplitude of the peaks, which has been ignored in the preceding analysis, it is doubtful whether one could assert with any degree of confidence which of the peaks were due to minor showers, and which peaks represented normal random

VOL. 14 1959

fluctuations in the background rate. Only the major showers can be identified with certainty.

This example underlines the statement, made earlier, that inspection of echo rates is an inadequate basis for recognition of weak shower activity.

4. A Statistical Test of Shower Activity

At least three investigations have been made of the probability distribution of the number of echoes detected in short time intervals (30 sec). Two of these (McCrosky, 1957; Bowden and Davies, 1957) agree that with these short intervals the distribution of meteor counts does not depart significantly from the Poisson distribution, and the latter authors advance reasons why the significant result found by the third (Wylie and Castillo, 1956) may not be real. Selected portions of the Adelaide record for December 1956 and January 1957 have been analysed in the same way, but using longer time intervals (2 min and 10 min). Once again no significant departures from the Poisson distribution have been found. Details will be found in Appendix I.

There is no physical reason, apart from shower activity, why clustering amongst meteors should occur over still longer intervals of time. Assuming then a random background activity, total counts within a prescribed interval of time should be distributed according to Poisson's law. The relative frequency of occurrence of a count n is given by $\exp(-m)m^n/n!$, in which the mean m can only be estimated from the meteor counts themselves. The probability of obtaining a count $\geqslant n$ is

$$P(n) = \sum_{x=n}^{\infty} \exp(-m)m^x/x\,! \tag{1}$$

By choosing an appropriate significance level S, those periods for which $P(n) < S$ may be identified. These are periods of suspected shower activity, which warrant further investigation by range-time plots of individual echoes. Similarly, consideration of the probability

$$P(n) = \sum_{x=0}^{n-1} \exp(-m)m^x/x\,! \tag{2}$$

in relation to the same significance level S indicates periods of abnormally low activity.

A basic time interval of $\frac{1}{2}$ hr has been chosen. This time interval is short in relation to the time of passage of most shower radiants through the collecting sector of the aerial.

Although the partial sums based on Poisson's law are exact, the probabilities (1) and (2) obtained from them are only approximate, because of the necessity to evaluate an expected count (E) from the meteor counts themselves, and then to accept this value of E as equivalent to the mean m. Even after excluding all known shower activity the expected count is by no means constant. It is influenced by both diurnal and seasonal variations in the background count. The seasonal change is so slow that over a short period, say a month, it may be ignored, but the rapid change in the diurnal pattern near midday in summer (cf. Fig. 2) and midnight in winter is a source of inaccuracy which cannot be eliminated. Another

factor which must be considered is short-term changes in the sensitivity of the equipment, due to minor faults, failure to achieve uniformity in film processing or slight interference by external noise. If severe, these faults are easily recognized and the affected portions of the record are rejected from the statistical analysis. These temporary and perhaps unrecognized changes in the threshold level of detection of echoes are much more serious than slow drifts in equipment sensitivity which may go unrecognized despite continual monitoring of the equipment.

The average diurnal variation for each month, after deletion of known shower activity and correction for all apparent short-term changes in sensitivity, has been

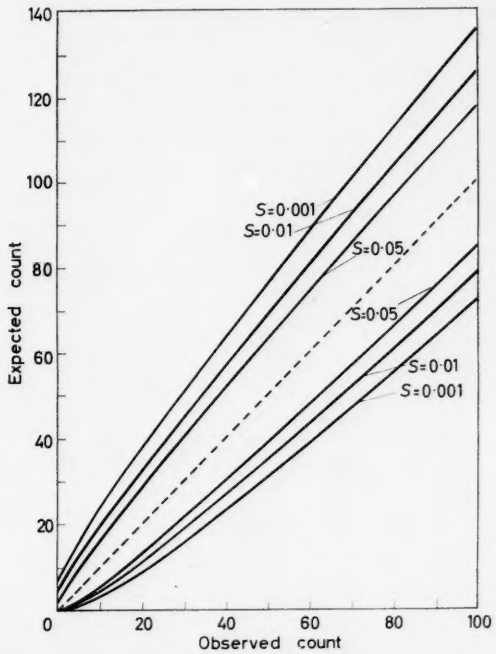


Fig. 4. Probability chart for estimating the significance level S of an observed meteor count.

accepted as the best estimate of the expected count (E) for each cell on any given day. In the computation of probabilities it is assumed that $E \equiv m$. The errors introduced into the E-values by the inclusion of cells which are subsequently identified as periods of shower activity can be eliminated by a process of iteration, if this is thought worth while.

In practice, nominal values of $S = 0{\cdot}05$, $0{\cdot}01$, $0{\cdot}001$ have been adopted. The type of chart in use is illustrated in Fig. 4. If this diagram is entered with the appropriate values of the observed count O and the expected count E, the level of significance to be attached to the particular count can be seen at a glance.

A similar search method for isolating shower activity has already been used by KRESÁKOVÁ and KRESÁK (1955) in their analysis of telescopic meteor observations in Czechoslovakia.

VOL. 14 1959

5. Examination of Meteor Counts for December 1956 and January 1957

The results of applying the statistical analysis outlined above to records covering 19 days in December 1956 and 21 days in January 1957 are illustrated in Fig. 5. This diagram shows the proportions of cells exhibiting significant departures from the mean behaviour for each ½ hr interval, the level of significance and the sense of the departure of the observed from the expected count. The known December showers—Geminids, Puppids, Phoenicids—have been excluded. No known January showers are observable from Adelaide.

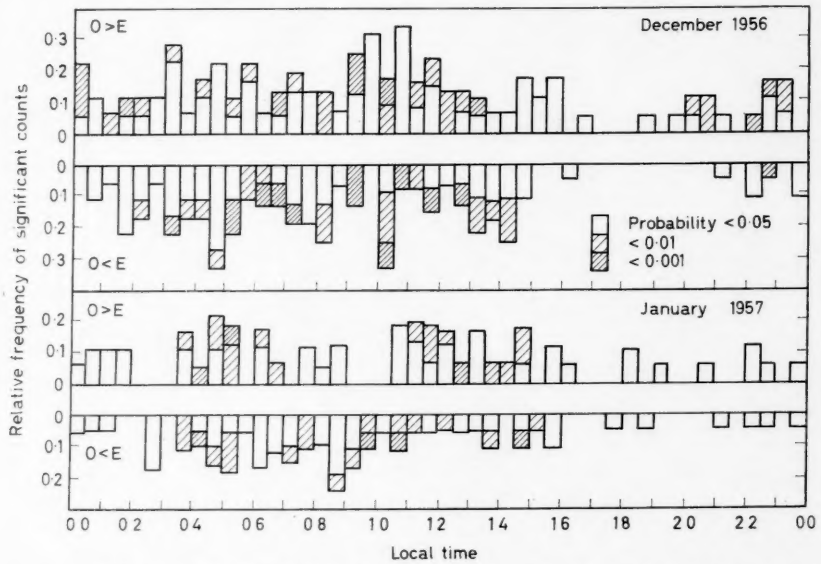


Fig. 5. Diurnal distribution of significant meteor counts, excluding showers.

It might be thought that at least some of the significantly low counts could be caused by interference by noise, but in fact the strongest and most persistent interference occurs just before 1800 hours LT, when the variability of the meteor counts is least. A detailed examination of the effect of noise interference shows that very few of the low counts can be attributed to this source. No attempt is made to read records contaminated by severe interference.

The most striking feature of these diagrams is the distribution, as a function of local time, of the cells whose counts differ significantly from the mean behaviour. The period of great variability, from shortly after midnight to shortly after noon, coincides with the time of passage, through the aerial collecting sectors, of the concentrations in the background radiant distribution. In contrast to this, the period near the time of passage of the antapex through the aerial collecting sectors is one in which the rate varies little from day to day. This surprising result is consistent with the observation by Hawkins (1956) that when meteors are observed in the hemisphere centred on the antapex there is a distribution of radiants more

uniform than at other times, when there is a strong concentration of radiants in the plane of the ecliptic.

Comparison of the number of significant cells observed with the number expected on a chance basis is made in Table 1. December shows a large excess of significant cells over the chance expectation. Although January records exhibit little more variability than expected, it must be remembered that the significant cells are not evenly distributed throughout the day, and that the difficulties in making correct allowance for short-term changes in sensitivity constitute an additional source of variability. A test of homogeneity of the hourly total counts for January, taking the records for each aerial separately, has also been made in the usual way. This test comes out very strongly against homogeneity from day to day. For the south aerial, $\chi^2 = 281$ with $n = 168$ degrees of freedom. Use of

Table 1. Comparison with expectation of observed numbers of significant meteor counts

$O - E$	Level of significance (%)	Month			
		December 1956		January 1957	
		Observed	Expected	Observed	Expected
Positive	5	92	41	59	46
	1	31	8	18	9
	0·1	11	1	3	1
Negative	5	81	42	59	46
	1	30	8	21	9
	0·1	12	1	4	1

the approximation $t = \sqrt{(2\chi^2)} - \sqrt{(2n - 1)}$ gives $t = 5{\cdot}42$ which is significant at the 10^{-7} level. For the north aerial, $\chi^2 = 383$, $n = 168$, $t = 9{\cdot}39$, significant at the 10^{-9} level. The contribution to χ^2 by the (hourly) cells between 1600 and 2200 hours is very small, consistent with Fig. 5. Shower activity precluded a similar homogeneity test for December.

A detailed comparison with partial rate plots shows, as expected, that only the most marked peaks and troughs in the partial rate plots depart significantly from the mean rate. Many of the significant counts, both high and low, arise in periods of some length—2 or 3 hr at least—when the whole level of meteor activity seems to rise above, or fall below, the mean level. Variation of this type is not suggestive of shower activity, and of course showers cannot be invoked in explanation of the periods of low count, which occur almost as frequently as high counts. This type of variability must rather be regarded as a property of the background activity, in whose study the statistical approach should prove fruitful.

A few shorter intervals of high count have been noted as periods of possible shower activity. These, together with data on the major showers, will be reported separately.

6. Conclusions

Statistical fluctuations in the background activity, the diffuseness of some of the shower radiants, and possibly the short time over which some showers are active, all impose limits to the recognition of weak shower activity in records obtained with narrow-beam radio equipments when these equipments are operated at echo rates practicable for a routine survey. Only the major showers, giving counts whose increases over the normal rate are highly significant, can be recognized with certainty and studied in detail; duration of the shower activity is then not a limitation to recognition. The remainder of the activity, after exclusion of recognized showers, is necessarily classed as background activity. Fortunately, the narrow collecting sector of the strongly-beamed aerial, operating on meteor radiants concentrated to the plane of the ecliptic, enables the distribution of the background activity to be studied in some detail, although there is always some ambiguity associated with an extended collecting area. Statistical studies of the pattern of the background activity then reveal significant variations from hour to hour and from day to day. Some of these variations are quite marked (although by no means as strong as the major showers), but they do not possess the structure associated with shower activity as detected by these equipments. This type of variability is tentatively identified as a characteristic of the true background activity. This is not to deny the existence of minor radiants; it merely emphasizes that most of the variability in meteor counts made with narrow-beam radio equipments is not of the type associated with recognizable showers in the sense defined in Section 2.

VOL.
14
195

Appendix I

Comparison of meteor counts with Poisson's distribution

This comparison has been made using selected portions of the meteor record for December 1956 and January 1957. If Poisson's Law is applicable, then as indicated in the text, the relative number of cells giving a count of n meteors is $\exp(-m)m^n/n!$ The mean meteor count per cell over the duration of the section of record examined is accepted as the mean, m, for the purpose of computation of probabilities.

Only those portions of the meteor record for which the average rate appeared by inspection to remain constant over a sufficiently long interval of time were selected for analysis. This inspection was made on either total or partial rate plots. Division into the basic short-duration cells was made in two ways: (a) 2 min in time, 200–800 km in slant range; (b) 10 min in time, 400–600 km in slant range. The chosen portions of the record are distributed throughout the day, and mean rates differ considerably. For this reason information obtained from different portions of the record has not been combined. No attempt has been made to use sections of the record yielding too few cells for application of the χ^2-test for goodness of fit. As range information is available to supplement time information there is no possibility of ambiguity in separation of echoes from different meteors, despite the slow film speed of 12 cm/hr.

Table 2 describes the records selected for analysis, and lists the probabilities found in the usual manner by the χ^2-test. Two typical observed distributions,

Table 2. Comparison of meteor counts with Poisson's Law

Type of analysis	Date	Time interval (hr)	Aerial	No. of cells	Mean cell count	Probability
Time:	Dec. 5–6	23–02	S	90	0·84	0·5
2 min	7	06–09	N	86	1·77	0·08
Range:	9	05–08	N	90	1·50	0·2
200–800 km	13	11–13	N	60	3·45	0·3
	15	01–05	N	120	0·89	0·5
	19	02–06	S	111	0·88	0·1
	20	07–11	N	120	1·23	0·97
	Jan. 3	06–10	N	105	3·30	0·9
	13–14	23·30–01·30	S	60	1·10	0·7
	15	01·30–03·30	S	60	1·20	0·2
	17	11–13	N	60	2·28	0·4
	18	03–06	N	88	1·66	0·9
	20	03–06	N	90	1·29	0·7
	27	02–08	S	180	0·72	0·4
	28	01–06	N	150	0·90	0·9
	31	02–04	S	60	1·88	0·05
Time:	Nov. 30–Dec. 1	22–05	N	43	2·56	0·2
10 min	Dec. 4–5	22–05	N	43	0·88	0·3
	6–7	22–09	N	65	2·40	0·5
Range:	8	03–10	S	43	1·84	0·2
400–600 km	8–9	22–04·30	N	40	1·45	0·6
	10	02–11	S	56	2·41	0·96
	10–11	22–05	N	43	1·30	0·4
	14–15	23–09	N	61	3·26	0·5
	16	02–09	S	46	1·96	0·2
	17	00–10	N	61	2·52	0·4
	20	05–11	N	37	3·08	0·01
	Jan. 20	01–10	N	54	3·83	0·95
	21–22	23–09	N	61	3·52	0·5
	27	02–09	S	44	2·52	0·93
	28	01–10	N	57	3·65	0·6
	28–29	15·30–00·30	S	55	0·55	0·5
	Dec. 2–3	10·20–10·30	N	146	3·26	10^{-80}
	12–13	11·30–13·00	N	154	3·29	10^{-54}

with the fitted Poisson distributions, are illustrated in Fig. 6. The last two lines of Table 2 give data for two records extending over more than 24 hr; one of these is also shown in Fig. 6. These latter examples were included to illustrate the effect of the diurnal variation, and to emphasize the importance of applying these tests only to data for which the mean count is known to be constant.

Of the thirty-two selected periods, only two show departures from the Poisson distribution significant at the 5 per cent level. This is no more than would be expected on the basis of chance events. These samples contain no evidence against the hypothesis that the background activity is randomly distributed.

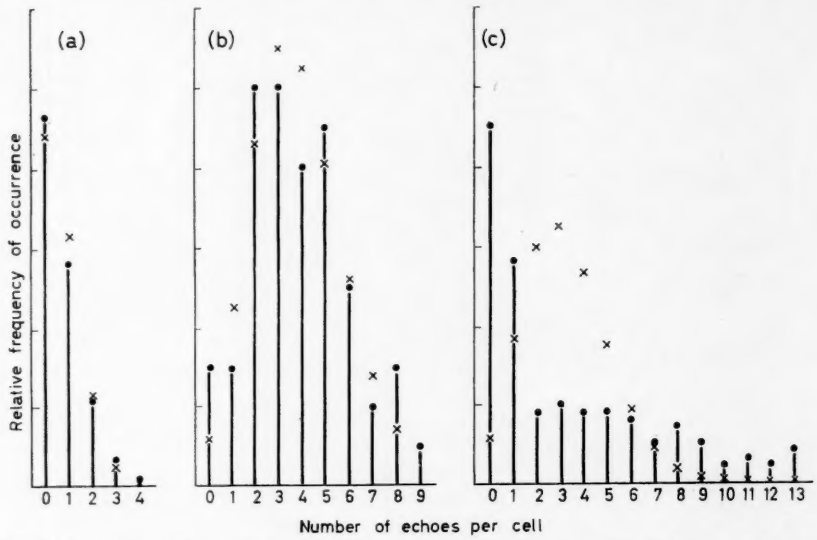


Fig. 6. Typical distributions in comparison of meteor echo counts with Poisson's distribution. ● Observed number of cells; × expected number of cells.
(A) 2 min cells; mean = 0·71; 6 hr record.
(B) 10 min cells; mean = 3·83; 10 hr record.
(C) 10 min cells; mean = 3·26; 24 hr record.

Acknowledgements—This work forms part of the meteor research programme of the University of Adelaide. It is financed jointly by the University of Adelaide, the Division of Radiophysics, C.S.I.R.O. and the Radio Research Board, C.S.I.R.O. The author has benefited greatly from discussions with Dr. C. D. Ellyett of Canterbury University College. The meteor survey equipment was operated by Mr. J. W. Smith.

References

Aspinall A., Clegg J. A., and Hawkins G. S.	1951	*Phil. Mag.* **42,** 504.
Bowden K. R. R. and Davies J. G.	1957	*J. Atmosph. Terr. Phys.* **11,** 62.
Clegg J. A.	1948	*Phil. Mag.* **39,** 577.
Gill J. C. and Davies J. G.	1956	*Mon. Not. R. Astr. Soc.* **116,** 105.
Hawkins G. S.	1956	*Mon. Not. R. Astr. Soc.* **116,** 92.
Keay C. S. L.	1957	*Aust. J. Phys.* **10,** 471.
Kresáková M. and Kresák L.	1955	*Contr. Astr. Obs., Skalnaté Pleso* **1,** 80.
McCrosky R. E.	1957	*Bull. Astr. Insts. Csl.* **8,** 1.
Weiss A. A.	1955	*Aust. J. Phys.* **8,** 148.
Weiss A. A.	1957	*Aust. J. Phys.* **10,** 77.
Wylie L. R. and Castillo H. T.	1956	*Ohio J. Sci.* **56,** 339.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 31 to 40. Pergamon Press Ltd. Printed in Northern Ireland

Geomagnetic influence on the *F1*- and *F2*-regions of the ionosphere—effect of solar activity

R. G. RASTOGI*
Physical Research Laboratory, Ahmedabad 9, India

(*Received* 13 *May* 1958)

Abstract—The paper discusses the geomagnetic influence on the noon critical frequencies of *F1*- and *F2*-layers in different seasons and at different stages of solar activity. The *F1*-layer is found to be affected by the magnetic field only in years of high solar activity. The equatorial depression of noon f_0F2 is, on the other hand, most prominent in years of low solar activity.

VOL. 14 1959

1. INTRODUCTION

IT is well known that the *F2*-layer of the ionosphere behaves anomalously in many ways, whereas the *E*- and *F1*-layers behave approximately as "Chapman layers". SEATON (1948) showed that the noon critical frequencies of *F1*-layer in January 1947 were not symmetrical about the subsolar point. The isopleths of average maximum electron concentration for the *F1*-layer indicated two principal high density regions at about 4°N and 21°S, while contours of the *E*-layer ionization were regularly disposed in agreement with contemporary ideas. LEJAY and ARDILLON (1950) indicated that the variation of the *F1*-layer differs somewhat from that of *E* and that it had some of the peculiarities of *F2*. GHOSH (1955) showed that the noon critical frequencies of *F1* were affected by the geomagnetic latitude in the same way as *F2*-layer critical frequencies and expressed the view that f_0F1 showed a depression at the magnetic equator both in high as well as in low sunspot years. EYFRIG (1955) supported these conclusions using f_0F1 data obtained by Japanese workers at a few equatorial stations in the minimum sunspot years 1943 and 1944. MAEDA (1955) and SATO (1955), however, did not find any trough in the latitudinal distribution of noon critical frequencies of *F1* during 1953. RASTOGI (1956, 1958) found that critical frequencies of *F1* extrapolated for zero sunspot number showed a definite maximum at the equator and that the occurrence of two maxima of f_0F1 in the middle latitudes was a feature of high sunspot periods. It may be mentioned that GHOSH had taken 1951 as a representative low sunspot years though the average Zürich sunspot number R_z in 1951 was 69·4. It was, therefore, felt desirable to re-examine the critical frequencies of *F1* and *F2* during years of different solar activity and this note gives some of the conclusions arrived at.

2. LATITUDINAL DISTRIBUTION OF NOON VALUES OF f_0E, f_0F1 AND f_0F2

First, the distributions of midday critical frequencies of *F1*-layer (mean of 11 to 13 hr) with respect to geographic latitude, geomagnetic latitude and magnetic dip are examined. We compare the distributions of f_0F1 with similar distributions of f_0E and f_0F2. In Figs. 1(a), 1(b) and 1(c) are plotted the midday critical

* Present address: Radio and Electrical Engineering Division, National Research Council, Ottawa, Canada.

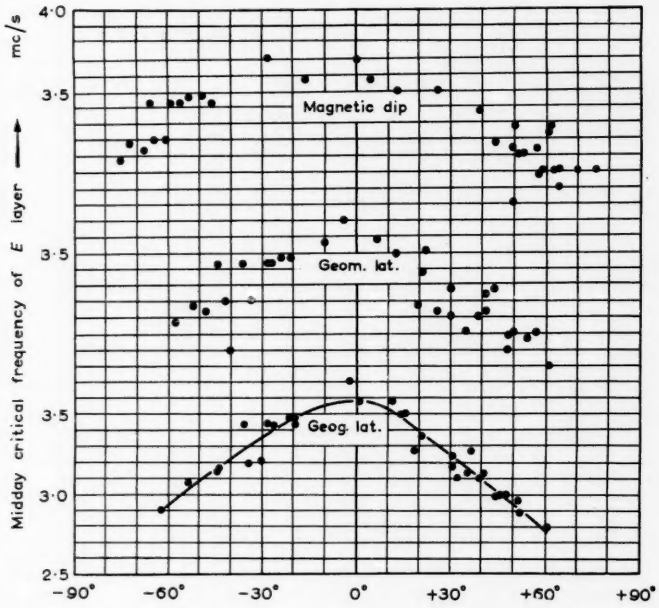


Fig. 1(a). Variation of midday (11–13 hr mean) critical frequencies of *E*-layer in March 1952 with geographic latitude, geomagnetic latitude and magnetic dip.

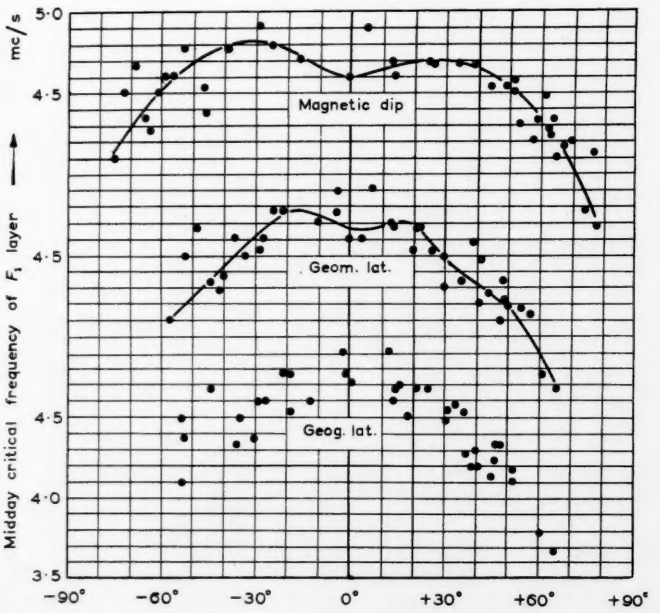


Fig. 1(b). Variation of midday (11–13 hr mean) critical frequencies of *F1*-layer in March 1952 with geographic latitude, geomagnetic latitude and magnetic dip.

frequencies of *E*, *F1* and *F2* in March 1952 against geographic and geomagnetic latitudes and magnetic dip.

The points relating to f_0E are least scattered in the plot against geographic latitude and the curve shows a maximum at the geographic equator. In the plots against geomagnetic latitude and magnetic dip, there is definitely greater scatter of points.

Referring to *F2* critical frequencies in Fig. 1(c)* the scatter of points is less in

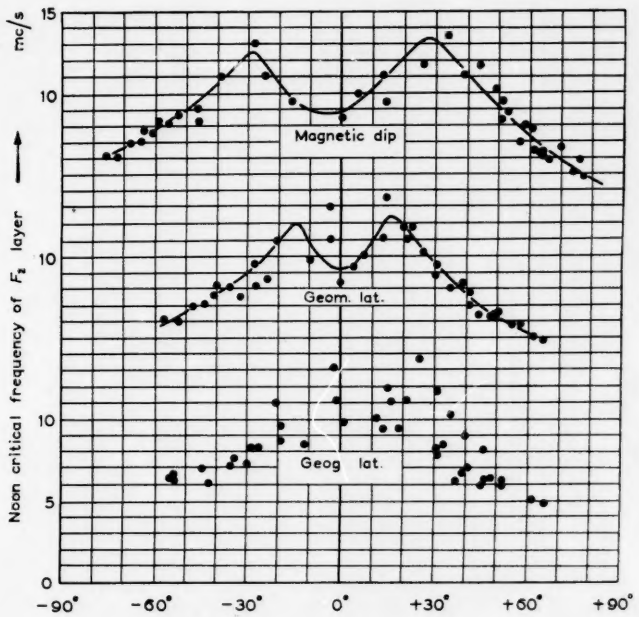


Fig. 1(c). Variation of noon critical frequencies of *F2*-layer in March 1952 with geographic latitude, geomagnetic latitude and magnetic dip.

the plots against geomagnetic latitude or magnetic dip and the curves show two maxima in middle latitudes (at $\pm 15°$ geomagnetic latitude or $\pm 30°$ magnetic dip) with a depression at the equator.

Referring to the plots of *F1* critical frequencies in Fig. 1(b), we find the largest scatter of points in the plot against geographic latitude and the least scatter in the plot against magnetic dip. Maximum frequencies are observed north and south of the equator at approximately the same latitudes as the f_0F2 peaks.

It may be concluded that the critical frequencies of *F2* and probably of *F1* are best arranged according to magnetic dip.

* It may be noted that in the plot of f_0F2 against geomagnetic latitude in Fig. 1(c), two points corresponding to Lwiro and Nairobi in Africa lie far from the smooth curve. On the plot against magnetic dip, however, these points fit very well on the smooth curve. In Fig. 1(b) also, the points corresponding to f_0F1 at Lwiro and Nairobi which are discrepant on the geomagnetic latitude plot, lie on a smooth curve when plotted against the magnetic dip. This discrepancy occurs in other months also at many African stations where there is large separation between the magnetic and geomagnetic equators.

VOL. 14 1959

3. Geomagnetic Control of Noon f_0F1 in Different Years within a Solar Cycle

Next, we examine the geomagnetic influence of f_0F1 at different stages of solar activity. In Figs. 2(a), 2(b) and 2(c) are shown the geomagnetic latitude distributions of noon critical frequencies of $F1$ averaged over (a) the equinoctial months, i.e. March, April, September and October (b) northern winter months, i.e.

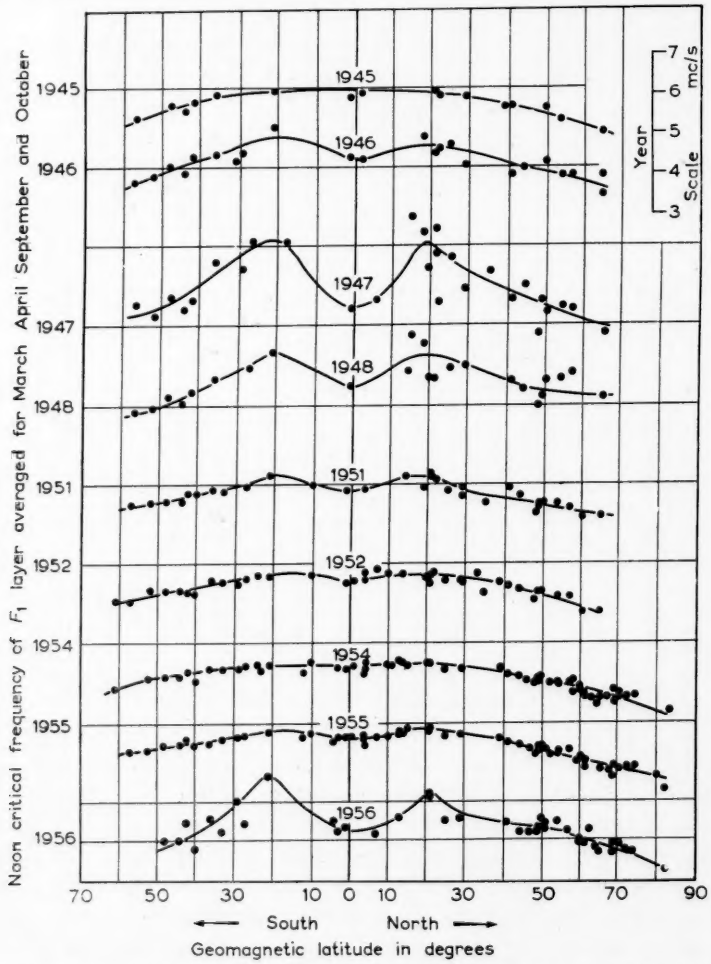


Fig. 2(a). Variations of noon critical frequencies of $F1$-layer with geomagnetic latitude for the equinoctial months (March, April, September and October) of individual years.

November, December, January and February; and (c) southern winter months, i.e. May, June, July and August, in a number of years from 1945 to 1956.

Referring to Fig. 2(a) it is seen that the depression of f_0F1 at the equator was most pronounced in the maximum sunspot years (1947 and 1948). The depression decreased steadily with decrease in solar activity in 1951 and 1952 and was absent

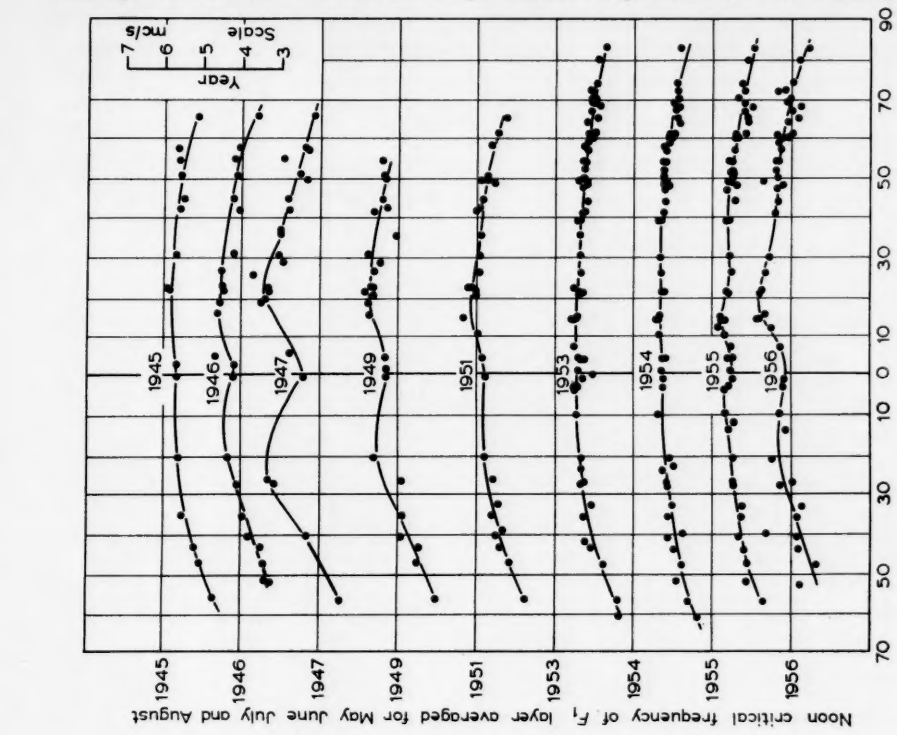


Fig. 2(c). Variations of noon critical frequencies of *F1*-layer with geomagnetic latitude for the southern solstitial months (November, December of preceding year and January, February of the following year).

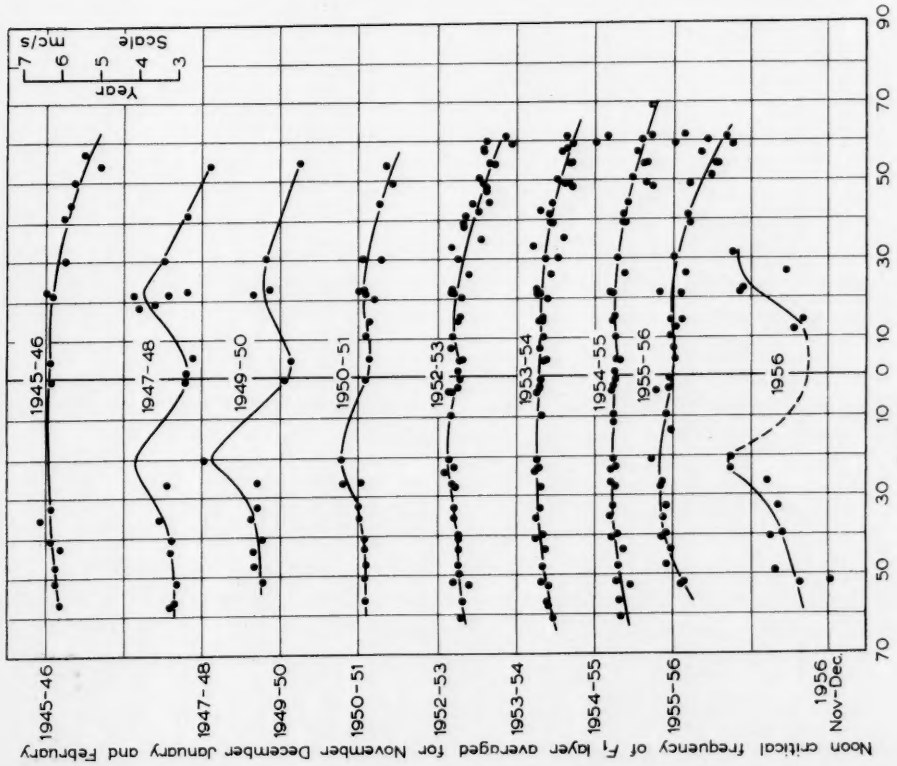


Fig. 2(b). Variations of noon critical frequencies of *F1*-layer with geomagnetic latitude for the northern solstitial months (May, June, July and August) of individual years.

VOL. 14 1959

in 1954, the year of least solar activity. It again developed in 1955 and was quite conspicuous in 1956, when the sunspot numbers had gone up. The two peaks occurred at about $\pm 20°$ geomagnetic latitude, and the north and south peaks were nearly equal in the equinoctial months.

Similar variations of equatorial depression of f_0F1 are observed in Figs. 2(b) and 2(c). The depressions are most conspicuous in the years of maximum solar activity (1947) and practically absent in a year of low solar activity (1954).

Table 1. Noon critical frequencies of *F1*- and *F2*-layers at the equatorial and middle latitude stations during equinoctial months of different years

Year	Zurich sunspot number	Noon critical frequency *F1*-layer				Noon critical frequency *F2*-layer			
		Subtropical Peak (Mc/s)	Equatorial trough (Mc/s)	Difference (Mc/s)	Ratio	Subtropical Peak (Mc/s)	Equatorial trough (Mc/s)	Difference (Mc/s)	Ratio
1945	39·3	4·85	4·85	0·0	1·00	(11·5)	(8·0)	(3·5)	(1·44)
1946	87·3	5·60	5·20	0·40	1·08	14·5	10·4	4·1	1·39
1947	153·2	7·20	5·50	1·70	1·31	15·2	13·0	2·2	1·17
1948	141·6	6·40	5·50	0·90	1·16	15·0	12·0	3·0	1·25
1949	145·4	6·50	5·40	1·10	1·21	15·4	12·4	3·0	1·24
1950	84·0	5·50	5·00	0·50	1·10	14·0	10·0	4·0	1·40
1951	70·9	5·20	4·90	0·30	1·06	13·7	9·7	4·0	1·41
1952	25·8	4·85	4·50	0·35	1·08	12·2	8·4	3·8	1·45
1953	18·7	4·60	4·45	0·15	1·03	11·4	7·5	3·9	1·52
1954	5·2	4·45	4·40	0·05	1·01	10·8	7·2	3·6	1·50
1955	28·7	4·85	4·60	0·25	1·05	11·7	8·0	3·7	1·46
1956	140·8	6·00	5·25	0·75	1·14	15·0	11·6	4·6	1·29
		Mean		0·54	1·10	Mean		3·5	1·38

We find that during equinoctial months the values of noon f_0F1 are fairly constant within $\pm 20°$ of the equator, beyond which it decreases with latitude. During solstitial months there is a flat maximum about the subsolar point.

In Table 1 are collected the values of f_0F1 and f_0F2 at the equatorial troughs and subtropical peaks during the equinoctial months of the years 1945 to 1956.

It will be noted that the ratio of f_0F1 at subtropical and equatorial regions was maximum in 1947 and 1956 and minimum in 1945 and 1954. It therefore appears that while there is no detectable geomagnetic influence on noon f_0F1 in minimum sunspot years, there is a pronounced effect in high sunspot years. The f_0F1 peaks are located at about $\pm 21°$ geomagnetic latitude, or $\pm 40°$ dip. The highest critical frequencies of *F1* were recorded in south hemisphere at Rarotonga (dip 39°) and in the north hemisphere at Okinawa (dip 37°) and Chungking (dip 43°).

4. Geomagnetic Control of Noon f_0F2 in Different Years within a Solar Cycle

In Fig. 3 are shown the variation with magnetic dip of noon f_0F2 in the equinoctial months of the individual years 1945 to 1956. The two peaks in f_0F2 occurred at magnetic dips of about $\pm 30°$. The southern peak seems to be smaller than the northern peak in most of the years, but as the number of stations near

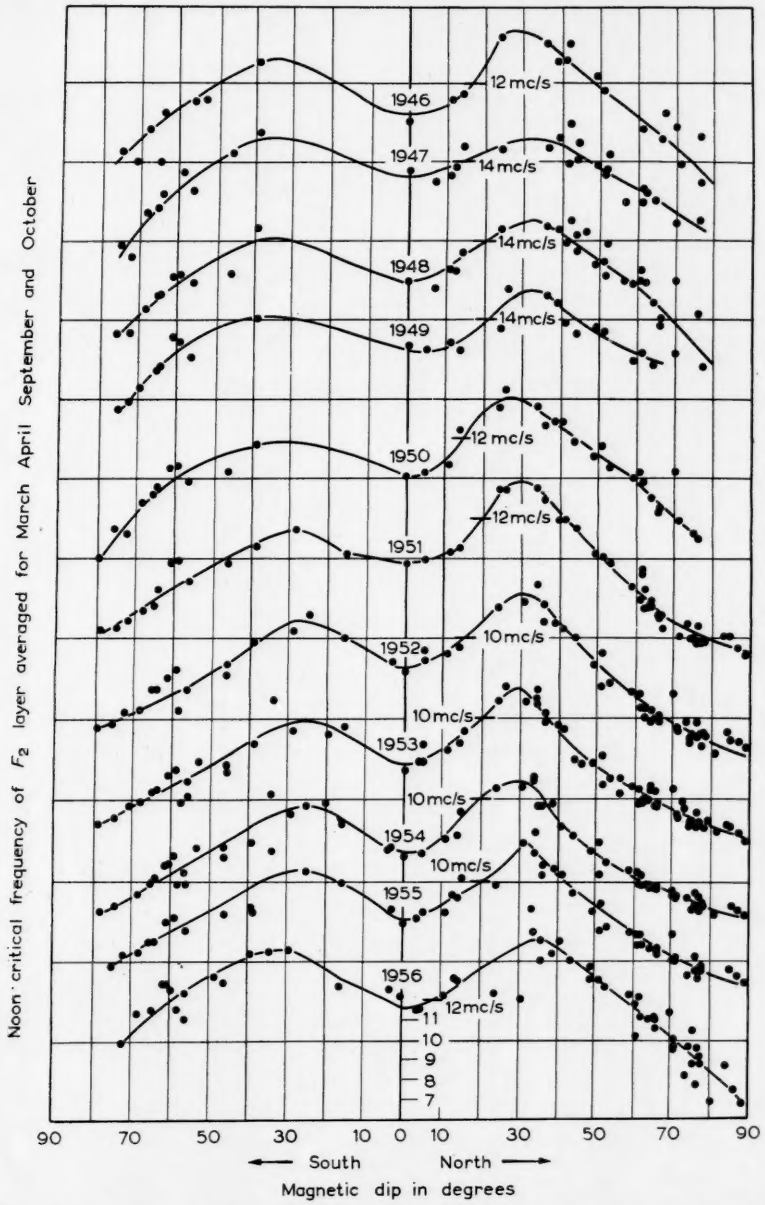


Fig. 3. Variations of noon critical frequencies of *F2*-layer with magnetic dip for the equinoctial months (March, April, September and October) of individual years.

the southern peak is smaller the conclusion should be taken with reserve. The difference between the northern peak and the equatorial trough values of f_0F2 in most of the years ranges between 3 and 4 Mc/s (refer to Table 1). The ratio of peak-to-trough values seems to be systematically higher during periods of lower

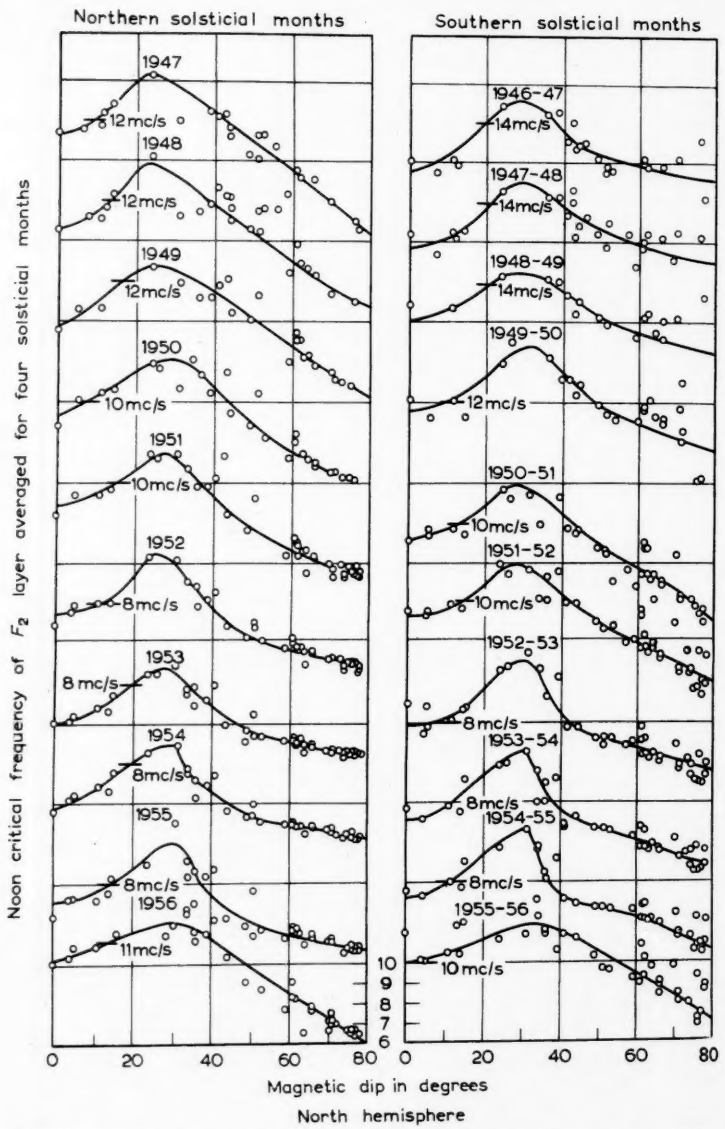


Fig. 4. Variations of noon critical frequencies of *F2*-layer in the north hemisphere with magnetic dip for the solstitial months (May, June, July and August) and (November, December, January and February) of individual years.

solar activity. Thus, the peak value of f_0F2 was 17 per cent greater than the trough value in the maximum sunspot year of 1947, while it was about 50 per cent greater in the minimum sunspot year of 1954.

In Fig. 4 are shown variations of noon f_0F2 with magnetic dip in the northern hemisphere during the solstitial months.

The relative equatorial depression of noon f_0F2 during sunspot minimum and

VOL
14
195

sunspot maximum periods can be clearly seen from Fig. 5, where are plotted the yearly average noon f_0F2 values in 1953 (low sunspot year) and 1956 (high sunspot year) against magnetic dip. Note that the f_0F2 scale is logarithmic and so the separation between the two curves at any latitude gives the proportional difference in f_0F2. It is evident from the figure that the equatorial depression was much reduced in 1956 compared to 1953.

It is to be noted that in 1953 (and also in 1954) the f_0F2 peak occurred at about 30° dip. The maximum f_0F2 (yearly average value) was observed at Calcutta

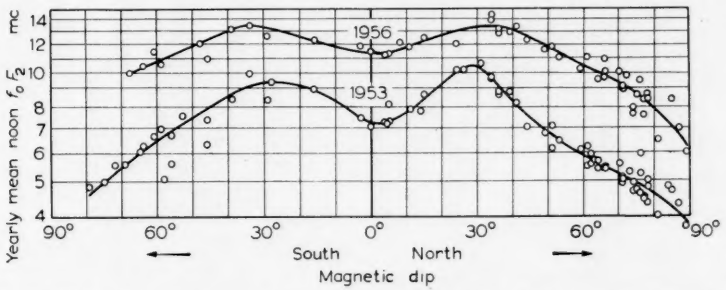


Fig. 5. Variations of yearly mean noon critical frequencies of *F2*-layer with magnetic dip in the years 1953 and 1956.

(dip 31°). In 1956, a year of large sunspot number, the f_0F2 peaks seem to have shifted away from the equator and were located at about 35° dip. The maximum f_0F2 (yearly average value) were observed at Formosa and Ahmedabad (dip 34°).

6. Conclusions

(1) During low sunspot years, noon critical frequencies of *F1* vary regularly with geographic latitude and the maximum value occurs near the sub-solar point.

(2) During high sunspot years, there develop two peaks of noon f_0F1 on either side of a magnetic equator with a depression at the equator.

(3) The latitudinal distribution of f_0F1 or f_0F2 shows less scatter when plotted against geomagnetic latitude or magnetic dip than against geographic latitude. The magnetic dip seems to be the most suitable co-ordinate for representing the variation of noon f_0F1 or f_0F2 with latitude.

(4) The peaks of f_0F2 occur at about 30° dip in low sunspot years and about 35° dip in high sunspot years. The peaks of f_0F1 occur farther from the equator at about 40° dip.

(5) The magnitude of equatorial depression of f_0F1 decreases progressively with decrease of sunspot number while that of f_0F2 remains fairly constant at different stages of solar activity. However, the proportionate equatorial depression of f_0F1 increases while that of f_0F2 decreases with increase in sunspot number.

Acknowledgement—The work described forms part of the programme of ionospheric research carried out in this laboratory under the guidance of Prof. K. R. Ramanathan, to whom the author is deeply indebted. The author is in receipt of financial assistance from the Council of Scientific and Industrial Research, India.

The data analysed were taken mainly from the *F*-Bulletins of Central Radio Propagation Laboratory, U.S.A.

REFERENCES

EYFRIG R.	1955	*C. R. Acad. Sci., Paris* A **241,** 759.
GHOSH M.	1955	*J. Geophys. Res.* **60,** 115.
LEJAY P. and ARDILLON Mlle.	1950	*Rec. Un. Radio-sci. Int.* Vol. VIII, Pt. II, p. 259.
MAEDA H.	1955	*Rep. Ionos. Res. Japan* **9,** 64.
RASTOGI R. G.	1956	*Studies in the Physics of the Ionosphere* Ph.D. Thesis, Gujarat University, Ahmedabad.
RASTOGI R. G.	1958	*Geofis. Pur. Appl.* **40,** No. 2, 145–155.
SATO T.	1955	*Rep. Ionos. Res. Japan* **9,** 205.
SEATON S. L.	1948	*J. Met.* **5,** 204.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 41 to 49. Pergamon Press Ltd. Printed in Northern Ireland

Ionospheric self-demodulation and self-distortion of radio waves

J. W. KING
Cavendish Laboratory, Free School Lane, Cambridge

(*Received* 16 *May* 1958)

Abstract—Experiments in which the phenomenon of ionospheric self-demodulation was investigated are described. Demodulation is observed to occur only at the lower modulation frequencies, and it is concluded that the occurrence of the phenomenon can be satisfactorily explained by an extension of BAILEY's theory of ionospheric cross-modulation as presented by HUXLEY and RATCLIFFE (1949). The magnitude of the observed demodulation effect agrees with the value predicted by theoretical considerations and is what would be expected in view of results obtained in cross-modulation experiments. Complications due to selective fading were not present in the experiments performed. found It is that self-demodulation decreases as does ordinary cross-modulation, near dawn; this is explained by the fact that the absorption of the radio wave occurs at a level where the collision frequency is of the same order of magnitude as the equivalent angular frequency of the wave.

VOL.
14
1959

1. INTRODUCTION

IT is well known that modulation can sometimes be transferred from one wave to another if both waves traverse an absorbing region of the ionosphere simultaneously, and the phenomenon has been called "wave-interaction", "Luxembourg effect" or "ionospheric cross-modulation". (BAILEY and MARTYN, 1934a, b; HUXLEY and RATCLIFFE, 1949). It is possible that a single modulated wave traversing an absorbing region in the ionosphere could produce cross-modulation on to itself and thus decrease its modulation to some extent.

It is the purpose of this paper to describe more fully the experiments reported by the author (KING, 1957) in which the phenomenon of ionospheric self-demodulation was investigated. The experiments are described in Section 3, and the main results are discussed in Section 4. In Section 5 it is shown that the amount of self-demodulation decreases as dawn approaches and this effect is discussed in relation to the well established, but hitherto unexplained, observation that the amount of cross-modulation also decreases with the approach of dawn.

Previous work (HUXLEY, 1952) on ionospheric cross-modulation has shown that the amount of transferred modulation does not always increase as much as would be expected at the lowest frequencies. In Section 6 this phenomenon is discussed and it is suggested that it arises because the disturbing wave suffers appreciable self-demodulation before its modulation is transferred to the wanted wave.

Other workers (AITCHISON and GOODWIN, 1955; CUTOLO, 1955) have observed a change in the coefficient of modulation of a wave caused by transmission through the ionosphere. Their results do not conform at all to the theory outlined in Section 2 and no explanation of them has been attempted in this paper. In the present series of experiments no phenomena of the type described by these other workers have been encountered.

HIBBERD (1955 and 1956) has published theoretical papers on the subject of self-demodulation and has arrived at conclusions similar to those presented in this paper. He has extended the theory (HIBBERD, 1957) to include effects which might occur near the gyro frequency.

2. THEORY OF SELF-DEMODULATION

BAILEY's theory of ionospheric cross-modulation, as summarized by HUXLEY and RATCLIFFE (1949), shows that when an unmodulated wanted wave traverses a region of the ionosphere which is also transmitting a disturbing wave modulated with a coefficient M at a frequency $\omega/2\pi$, its amplitude is modulated by a factor m_c given by

$$m_c = 1 - T_\omega \cos(\omega t - \phi_1) - T_{2\omega} \cos(2\omega t - \phi_2) \tag{1}$$

where

$$T_\omega = T_0[1 + (\omega/Gv)^2]^{-\frac{1}{2}} \tag{2}$$

$$T_{2\omega} = \tfrac{1}{4} M T_0[1 + (2\omega/Gv)^2]^{-\frac{1}{2}} \tag{3}$$

$$\tan \phi_1 = \tfrac{1}{2} \tan \phi_2 = \omega/Gv$$

and T_0 is the value to which T_ω would tend as the modulation frequency ω

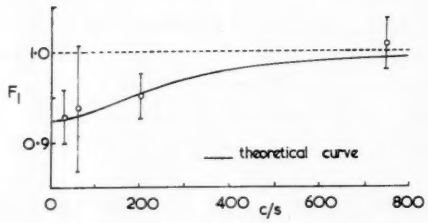


Fig. 1. Theoretical and experimental values of F_1. The curve is drawn for the case $M = 0{\cdot}3$ and $T_0 = 0{\cdot}024$.

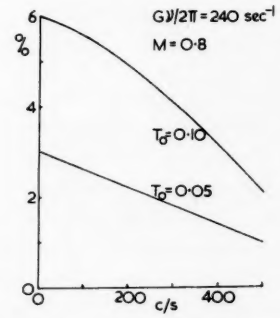


Fig. 2. Theoretical curves showing the percentage of second harmonic introduced.

becomes small. T_0 depends on the power and the modulation depth of the disturbing wave, and on the radio frequencies of both waves. T_0 can be determined for any pair of senders by making measurements on a series of modulation frequencies and using equation (2), and it has thus been shown that it is proportional to the power, and to the coefficient of modulation, of the disturbing wave as would be expected from theory.

If now a single wave modulated so that its amplitude varies as

$$1 + M \cos \omega t$$

traverses the ionosphere we shall suppose that it modifies its own amplitude by the factor m_c given in equation (1) so that it emerges with its amplitude modulated by a factor m_d where

$$m_d = (1 + M \cos \omega t) m_c \tag{4}$$

It is shown in the Appendix that equation (4) may be written

$$m_d = 1 + MF_1 \cos(\omega t + \alpha) + MF_2 \cos(2\omega t + \beta) \tag{5}$$

where $$F_1 = 1 - (\tfrac{1}{2}T_0/M^2)(2M - M^3 - T_0)\cos^2\phi_1 - \tfrac{1}{8}MT_0\cos^2\phi_2 \quad (6)$$

$$F_2 = \frac{\tfrac{1}{2}T_0[\tfrac{1}{4}\cos^2\phi_2 + \cos^2\phi_1 + \cos^2\phi_1\cos^2\phi_2 + \cos\phi_1\sin\phi_1\cos\phi_2\sin\phi_2]^{\frac{1}{2}}}{1 - \tfrac{1}{2}MT_0\cos^2\phi_1} \quad (7)$$

and α and β are small phase advancements.

The factor F_1 given by equation (6) represents the demodulation produced by transmission through the ionosphere; its variation with frequency is shown in Fig. 1 for a case where $M = 0{\cdot}3$ and $T_0 = 0{\cdot}024$. The quantity MF_2, representing the amplitude of the second harmonic introduced, is plotted in Fig. 2 for two values of T_0.

3. Experimental Arrangements

Many experiments made in the past at Cambridge have made use of the BBC sender at Droitwich (200 kc/s, 200 kW) to provide the disturbing wave, and measurements of cross-modulation have shown that $T_0 \simeq 0{\cdot}04$ when the modulation coefficient $M = 1$. The wanted wave in these experiments was provided by different senders on different occasions. Because so much was known about cross-modulation under these conditions it was decided to investigate self-modulation on a wave emitted from Droitwich, and to see whether the effect was consistent with that expected from the magnitude of T_0 observed on cross-modulation. The measurements were made at night at Cambridge, distant 150 km from Droitwich.

VOL. 14 1959

In order to measure any small change in modulation depth the coefficients of modulation on the ground wave alone and then on the sky wave alone were measured with the same equipment used under the same circumstances, and the mean results compared. It was known from other measurements that the midday field strength at the receiving point was almost entirely due to ground wave alone. The sky wave could therefore be isolated by using a loop aerial adjusted, at midday, so that no detectable signal was received. At night an e.m.f. representing the ground wave was obtained by receiving on a vertical aerial. From observations of the fading of the signal received on such an aerial it was concluded that the sky-wave component was of negligible magnitude.

Fading caused the amplitude of the sky wave to vary over a wide range and it was important that this fading should not introduce spurious results. The amplifier was provided with automatic gain control so that the mean radiofrequency e.m.f. applied to the measuring device was always the same. The magnitude of this e.m.f. was read on a meter and frequently checked for constancy. The depth of modulation was measured with an apparatus described by Shaw (1951). In this apparatus the constant radiofrequency e.m.f. was demodulated and the audiofrequency e.m.f. was applied to circuits which produced on a cathode-ray oscillograph a vector which indicated the amplitude and phase of the modulation. The oscillograph screen was photographed and the records later measured up. The phase of the modulation remained constant to within 1–2° during most of the fading. This fact provides support for the conclusion reached below that selective fading was not important. On a very few occasions the phase was observed

to be irregular; it was supposed that selective fading was then important and the corresponding measurements were rejected.

If, during the fading, the amplitude of the carrier and side bands change by different amounts the received depth of modulation will change. The fading is caused by an interference mechanism, and if there is anywhere in the transmission system a path difference of length d then two waves whose frequencies differ by f will not fade appreciably differently if $f < c/4d$ where c is the free-space velocity. If then f is the modulation frequency there should be fairly complete freedom from selective fading of this kind if $d < \frac{1}{8}c/f$. In the experiments to be described $f < 800$ c/s so that spurious results would be caused by fading if d is greater than about 50 km. The waves which were used were normally reflected from the E-layer and the reflection coefficient was small enough for the second reflection to be unimportant on most occasions. Path differences of the order of 50 km were therefore not expected. The most important measurements were made with modulation frequencies of 200 c/s or less, for which the path difference would have to be >200 km for selective fading to be important.

In this connexion it should be noted that in experiments on self-demodulation performed by other workers (AITCHISON and GOODWIN, 1955; CUTOLO, 1955) the radiofrequencies used were probably reflected from the F-layer, and possibly from the E-layer simultaneously, and the modulation frequencies were considerably greater. Under these conditions selective fading must presumably have been important, and it must have been difficult to allow for its effects.

VOL
14
195

4. RESULTS OF EXPERIMENTS ON SELF-DEMODULATION

4.1. *Dependence on modulation frequency*

Experiments were made on eleven different nights. On most of these the Droitwich sender (200 kc/s, distant 150 km) was used, but on three nights a sender at Daventry (647 kc/s, distant 90 km) was used. On most occasions modulation frequencies of 50, 100, 200 and 750 c/s were used. The procedure adopted was that, for each frequency, observations were made for 3 min on the ground wave and 12 min on the sky wave, measurements being taken at the rate of 20/min. Results with all four modulation frequencies were thus obtained in 1 hr, and the process could be repeated several times during a night. The power and the modulation depth were varied on different occasions.

On all occasions the demodulation, measured by the factor F_1 in equation 5, was found to be more pronounced at the lower frequencies. The results for a series of three nights on which the power was 200 kW and the coefficient of modulation 0·3 are shown in Fig. 1 where the lengths of the lines indicate the standard deviations of the readings. Each point corresponds to about 1500 separate determinations of the ratio of the modulation coefficients.

4.2. *Dependence on power and modulation depth*

Experiments made with different conditions at the sender gave similar results; for each set of conditions curves showing F_1 as a function of modulation frequency were plotted and the best fitting curve used to give a value of T_0. Theoretical and

experimental investigations of cross-modulation have shown that T_0 is proportional to the power of the sender and to the coefficient of modulation, so that it is possible to calculate what T_0 would have been under standard conditions of 100 kW with 100 per cent modulation. The values of T_0 normalized in this way are shown in Table 1. The final results, obtained using a wide range of transmitted power, are reasonably consistent.

TABLE 1. Measured values of T_0 normalized to a power of 100 kW and to $M = 1$

Dates	Frequency (kc/s)	Power (kW)	M	T_0 normalized
27–29/1/1954	647	75	0·2	0·034
18–21/5/1954	200	400	0·2	0·037
22 and 24/6/1955 and 5/7/1955	200	200	0·3	0·040
7/7/1955	200	400	0·3	0·037

VOL. 14 1959

4.3. *The magnitude of the effect*

From Table 1 it appears that the mean value of T_0 (normalized for 100 kW and 100 per cent modulation) is about 0·037. It is interesting to compare this experimental value with estimates made in other ways:

(a) T_0 may be calculated from the theory of cross-modulation as given by HUXLEY and RATCLIFFE (1949). Their equation 15 reads

$$T_0/\alpha = Me^2E_D^2/(mp_D^2 3Gk\theta) \tag{8}$$

where, in the present case, α represents the attenuation in nepers of the wave itself; p_D is the effective angular radiofrequency of the wave which for an actual frequency of 200 kc/s is approximately equal to the angular gyrofrequency $p_H = 2\pi \times 1{\cdot}25 \times 10^6$ c/s; E_D is the electric field strength of the wave $= 8 \times 10^{-7}$ e.s.u. for 100 kW radiated at an angle of 45° to a height of 100 km; G is assumed to be 2×10^{-3}; e and m are the charge and mass respectively of the electron; k is Boltzmann's constant; θ, the temperature, is taken to be 193°K, i.e. the temperature given by NICOLET (1952) for a height of 90 km assuming oxygen half-dissociated.

Insertion of the above quantities into equation 8 shows that for $M = 1{\cdot}0$ we obtain $T_0/\alpha = 0{\cdot}0161$, i.e. for 1 neper of absorption T_0 is 1·61 per cent. The best possible estimate (ALLCOCK, 1954; BELROSE, 1955) of the absorption suffered by a wave from Droitwich received in Cambridge is 2·21 nepers, and therefore the theoretical estimate of T_0 for the self-demodulation which would be expected in the experiments performed is $2{\cdot}21 \times 1{\cdot}61$ per cent $= 0{\cdot}036$. This value is in good agreement with the experimental value of T_0 deduced from self-demodulation experiments.

(b) We may compare the magnitude of T_0 deduced from self-demodulation

experiments with that deduced from cross-modulation experiments. The theory given by HUXLEY and RATCLIFFE can be extended to show that if T_{0C} and T_{0D} are the magnitudes of T_0 prevailing in cross-modulation and self-demodulation experiments respectively then $T_{0C}/T_{0D} = \alpha_W/\alpha_D$ where α_W and α_D represent the absorbtion of the wanted wave and the single wave in the two cases respectively. A careful study of all the Cambridge cross-modulation experiments indicates that for the standard case of 100 kW of disturbing power and 100 per cent transmitted modulation the quantity $T_{0C}/\alpha_W = 0{\cdot}013$/neper. This quantity compares well with the value for $T_{0D}/\alpha_D = 0{\cdot}037/2{\cdot}2 = 0{\cdot}017$/neper observed in self-demodulation experiments. The fact that the observed value of T_{0D}/α_D is slightly larger than T_{0C}/α_C might be expected since in cross-modulation experiments the absorbing regions for the wanted and disturbing waves might not always overlap completely.

4.4. *The production of harmonics*

Observations were made on two nights with a wave radiated from Droitwich (200 kc/s, 200 kW) and modulated at 200 c/s with $M = 0{\cdot}6$. A Marconi wave analyser was used to measure the amount of second harmonic present. In order to measure the amount of harmonic content introduced by the process of self-distortion the amount of second harmonic present on the ground wave had to be measured to allow for any harmonic content introduced by the apparatus. More than 400 measurements were made on the sky wave and 100 on the ground wave, and it was found that the process of self-distortion introduced 1·5 per cent of second harmonic into the sky wave.

A reasonable estimate for the quantity $G\nu/2\pi$ for the region where the self-distortion occurred during the experiments is $G\nu/2\pi = 350$ sec^{-1}, and upon substituting this value into equatoin 7 the value of T_0 (normalized to 100 kW and $M = 1$) is found to be 0·035, which is in good agreement with the results given above.

5. PHENOMENA OCCURRING NEAR DAWN

It is well established that the magnitude of ionospheric cross-modulation decreases during the hour before sunrise and is very small at sunrise (RATCLIFFE and SHAW, 1948). One possible explanation (SHAW, 1951) is that whereas during the night the appropriate value of the collision frequency ν is much smaller than the equivalent angular frequency p (given by $p = p_D + p_L$ for self-demodulation and $p = p_W + p_L$ for cross-modulation, where p_L is the angular gyrofrequency) it cannot be supposed that ν remains small at the levels where the wave is absorbed near dawn. When ν and p are of the same order of magnitude the magnitude of T_0 will not depend on $1/p^2$, but it will be proportional to

$$(p^2 - \nu^2)/(p^2 + \nu^2)^2 \qquad (9)$$

An alternative explanation (RATCLIFFE and SHAW, 1948) supposes that the regions where the wanted and the disturbing waves are absorbed overlap during the night but are separate near dawn.

The present experiments throw some light on this phenomenon because they

VOL
14
195

show that the amount of self-demodulation also decreases as dawn approaches. Since this observation cannot be explained in terms of the separation of the absorbing regions it favours the first explanation given above.

An attempt was made to relate the magnitude of v, deduced from cross-modulation experiments, to the magnitude of T_0 observed in the same experiments as dawn approached. Droitwich (200 kc/s, 200 kW) was used as the disturbing sender and Westerglen (809 kc/s, 100 kW, distant 450 km) provided the wanted wave. The quantity $Gv/2\pi$ was determined by both the well known phase and

Table 2. Values of $Gv/2\pi$ and the corresponding transferred modulation obtained on 3/6/1954

GMT	Transferred modulation (%)	$Gv/2\pi$ ($\sec^{-1}$)
Night-time	3·1	230
0240–0300	2·6	440
0300–0320	1·1	1000

amplitude methods, and the two methods gave concordant results, a typical set of which is given in Table 2. The corresponding values of the transferred modulation are also given.

For the conditions of the experiment calculation shows that the results given in Table 2 agree with equation 9 if the constant G is taken as $2{\cdot}0 \times 10^{-3}$.

6. The "Low Frequency Anomaly" in Cross-Modulation

It has been found (Huxley, 1952) that although equation (2) represents the relation between T_ω and ω, for frequencies greater than about 250 c/s, it often gives values of T_ω which are greater than those actually observed at the lower modulation frequencies. In the observed results there is, indeed, often found to be a maximum of transferred modulation at some frequency in the region of 150 c/s, such that the value of T_ω is the same for 60 c/s as for a frequency of about 250 c/s.

A possible explanation of this phenomenon could be given by supposing that at the lowest modulation frequencies the disturbing wave suffered appreciable self-demodulation before it had produced the greater part of the cross-modulation. Some of the consequences of this explanation may be tested as follows:

A reasonable estimate of the magnitude of the low frequency anomaly is that the transferred modulation at 60 c/s is about 0·9 times what it ought to be on simple theory. If the suggested mechanism were correct, and if all the self-demodulation occurred before cross-modulation was produced, this implies that $F_1 = 0{\cdot}9$, and if we assume that the experiments were made with a disturbing power of 200 kW and with $M = 0{\cdot}6$, we conclude that we would require a value of T_0 (normalized) of 0·062. However, the value expected for T_0 is 0·037, which is too small by a factor of about two, but is nevertheless of the right order of magnitude.

A detailed examination of the phenomenon reveals two other facts which would be accounted for on the explanation here suggested. The first of these is that the low frequency anomaly becomes smaller with the onset of dawn, just as the self-demodulation becomes smaller. The second is that at night whenever the cross-modulation is found to be unusually large the low frequency anomaly is more pronounced.

APPENDIX

We have equation (4):

$$\begin{aligned} m_d &= (1 + M\cos\omega t)[1 - T_\omega\cos(\omega t - \phi_1) - T_{2\omega}\cos(2\omega t - \phi_2)] \\ &= 1 + M\cos\omega t - T_\omega\cos(\omega t - \phi_1) - T_{2\omega}\cos(2\omega t - \phi_2) \\ &\quad - MT_\omega\cos\omega t\cos(\omega t - \phi_1) - MT_{2\omega}\cos\omega t\cos(2\omega t - \phi_2) \\ &= 1 + \cos\omega t[M - T_\omega\cos\phi_1 - \tfrac{1}{2}MT_{2\omega}\cos\phi_2] \\ &\quad + \sin\omega t[-T_\omega\sin\phi_1 - \tfrac{1}{2}MT_{2\omega}\sin\phi_2] \\ &\quad + \cos 2\omega t[-T_{2\omega}\cos\phi_2 - \tfrac{1}{2}MT_\omega\cos\phi_1] \\ &\quad + \sin 2\omega t[-T_{2\omega}\sin\phi_2 - \tfrac{1}{2}MT_\omega\sin\phi_1] \\ &\quad - \tfrac{1}{2}MT_{2\omega}\cos(3\omega t - \phi_2) - \tfrac{1}{2}MT_\omega\cos\phi_1 \qquad \text{(A1)} \\ &= 1 + A\cos(\omega t + \alpha) + B\cos(2\omega t + \beta) + C\cos(3\omega t + \gamma) + D \qquad \text{(A2)} \end{aligned}$$

where α, β and γ are small phase advancements which may be readily calculated if required. D is seen to be given by $-\frac{1}{2}MT_\omega\cos\phi_1$, and

$$A^2 = [M - T_\omega\cos\phi_1 - \tfrac{1}{2}MT_{2\omega}\cos\phi_2]^2 + [T_\omega\sin\phi_1 + \tfrac{1}{2}MT_{2\omega}\sin\phi_2]^2$$

Neglecting terms containing $(T_{2\omega})^2$ and $(T_{2\omega})(T_\omega)$ we find:

$$A^2 = M^2 + {T_\omega}^2 - 2MT_\omega\cos\phi_1 - M^2T_{2\omega}\cos\phi_2.$$

Hence, using equations (2) and (3), we have:

$$A^2 = M^2[1 + ({T_0}^2/M^2)\cos^2\phi_1 - (2T_0/M)\cos^2\phi_1 - \tfrac{1}{4}MT_0\cos^2\phi_2]$$

therefore $$A = M\{1 - (T_0/M)[1 - (\tfrac{1}{2}T_0/M)]\cos^2\phi_1 - \tfrac{1}{8}MT_0\cos^2\phi_2\}$$

Now we can re-write equation (A2) in the form:

$$m_d = 1 + MF_1\cos(\omega t + \alpha) + MF_2\cos(2\omega t + \beta)$$

(neglecting the term representing the third harmonic) where

$$MF_1 = \frac{A}{1 + D} \quad \text{and} \quad MF_2 = \frac{B}{1 + D}$$

therefore $$F_1 = \frac{1 - (T_0/M)[1 - (\tfrac{1}{2}T_0/M)\cos^2\phi_1 - \tfrac{1}{8}MT_0\cos^2\phi_2}{1 - \tfrac{1}{2}MT_0\cos^2\phi_1}$$

$$= 1 - (\tfrac{1}{2}T_0/M^2)(2M - T_0 - M^3)\cos^2\phi_1 - \tfrac{1}{8}MT_0\cos^2\phi_2$$

VOL
14
195

Similarly we see from equation (A1) that

$$B^2 = T_{2\omega}{}^2 + \tfrac{1}{4}M^2T_{2\omega}{}^2 + T_\omega T_{2\omega} M \cos(\phi_1 - \phi_2)$$
$$= \tfrac{1}{4}M^2T_0{}^2[\tfrac{1}{4}\cos^2\phi_2 + \cos^2\phi_1 + \cos\phi_1\cos\phi_2\cos(\phi_1 - \phi_2)]$$

therefore

$$F_2 = \frac{\tfrac{1}{2}T_0[\tfrac{1}{4}\cos^2\phi_2 + \cos^2\phi_1 + \cos^2\phi_1\cos^2\phi_2 + \cos\phi_1\sin\phi_1\cos\phi_2\sin\phi_2]^{\frac{1}{2}}}{1 - \tfrac{1}{2}MT_0\cos^2\phi_1}$$

Acknowledgements—The author is grateful to Mr. J. A. RATCLIFFE for his encouragement, and to the B.B.C. for providing the special transmissions for the experiments described.

REFERENCES

AITCHISON G. J. and GOODWIN G. L.	1955	*Nuovo Cim.* **1,** 722.
ALLCOCK G. McK.	1955	*Ionosphere Conference Report* p. 14. Physical Society, London.
BAILEY V. A. and MARTYN D. F.	1934a	*Nature, Lond.* **133,** 218.
	1934b	*Phil. Mag.* **18,** 369.
BELROSE J. S.	1955	Private communication.
CUTOLO M.	1955	*Nuovo Cim.* **1,** 726.
HIBBERD F. H.	1955	*J. Atmosph. Terr. Phys.* **6,** 268.
HIBBERD F. H.	1956	*J. Atmosph. Terr. Phys.* **8,** 121.
HIBBERD F. H.	1957	*J. Atmosph. Terr. Phys.* **11,** 102.
HUXLEY L. G. H.	1952	*Nuovo Cim.* **9,** Suppl. p. 1.
HUXLEY L. G. H. and RATCLIFFE J. A.	1949	*Proc. Instn. Elect. Engrs.* Pt. III, **96,** 433.
KING J. W.	1957	*J. Atmosph. Terr. Phys.* **10,** 166.
NICOLET M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
RATCLIFFE J. A. and SHAW I. J.	1948	*Proc. Roy. Soc.* A **193,** 311.
SHAW I. J.	1951	*Proc. Phys. Soc. Lond.* B **64,** 1.

VOL. 14 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 50 to 62. Pergamon Press Ltd. Printed in Northern Ireland

Ray paths in the ionosphere

Approximate calculations in the presence of the earth's magnetic field

J. E. TITHERIDGE*
Auckland University College, New Zealand

(*Received* 16 *May* 1958)

Abstract—Expressions are derived from which it is possible to plot approximate ray paths through "linear" or "parabolic" ionospheric layers in the presence of the earth's magnetic field. From the curves of Figs. 3, 4 and 5 it is possible to compute very simply the horizontal displacement of the reflection point for any dip angle, and for vertical or oblique incidence.

1. INTRODUCTION

WITH the inclusion of the earth's magnetic field the equations governing the propagation of radio waves in the ionosphere become very complex, and exact analytical solutions to many of the problems encountered are difficult and often impossible in terms of simple functions. BOOKER (1938 and 1949) showed that the magnetic field causes deviations of a ray out of the plane of incidence, and gave a graphical method of tracing the ray paths in certain cases. POEVERLEIN (1950) constructed some paths for the two-dimensional case when the direction of the magnetic field lies in the plane of incidence. He used a simpler graphical method and indicated how the paths might be expected to change under different conditions. FORSGREN (1951) calculated some vertical-incidence paths for conditions in Sweden. Also, by a rough approximation (considering only the larger of the two terms under the square root in equation (1)) he showed how the deviation at vertical incidence may change with varying frequency and magnetic field. A few workers have since published ray paths derived by POEVERLEIN'S method for particular conditions. The work involved is considerable, and any graphical method is subject to accumulated errors. Some exact ray paths at oblique incidence have recently been calculated by HASELGROVE (1957) using an electronic computer. Curves are given showing the effect of the magnetic field on m.u.f. calculations for propagation in the magnetic meridian. The present paper attempts to derive general formulae for rapidly determining the approximate ray paths and lateral deviations of high frequency radio waves under a wide range of conditions. The curvature of the ionosphere has very little effect on the deviations from the no-field path and will be ignored.

2. SIMPLIFICATION OF FORMULAE

For conditions in the upper ionosphere (when the effect of collisions and the Lorentz polarization term may be ignored) the Appleton–Hartree equation for the normal refractive index n may be written

$$\frac{X}{1-n^2} = 1 + Y\cos\theta\,\{-q \pm \sqrt{(1+q^2)}\} \qquad (1)$$

* Now at the Cavendish Laboratory, Cambridge.

where $$X = \frac{Ne^2}{\pi m f^2}, \qquad Y = \frac{f_H}{f}, \qquad q = \frac{Y \sin^2 \theta}{2(1 - X)\cos \theta}$$

and N is the electron density, f the wave frequency, f_H the gyrofrequency and θ the acute angle between the directions of the wave-normal and the magnetic field. The upper and lower alternative signs refer to the Ordinary and Extraordinary rays respectively.

Simpler forms of this relation may be derived from the approximation

$$\sqrt{(1 + q^2)} = q + \frac{1}{1 + q} + \xi \tag{2}$$

where ξ is positive and less than the least of $1/10$, $q/10$, and $(1 + q)/24$ for all values of q. We thus obtain, for the Ordinary ray

$$\frac{X}{1 - n^2} = 1 + \frac{Y \cos \theta}{1 + q} \text{ with a } +\text{ve error} < \frac{Y}{10} \cos \theta, \frac{Y^2 \sin^2 \theta}{20(1 - X)} \tag{3}$$

$$\frac{X}{1 - n^2} = 1 + Y \cos \theta \text{ with a } +\text{ve error} < \frac{Y^2 \sin^2 \theta}{2(1 - X)(1 + q)} \tag{4}$$

$$n^2 = 1 - X \text{ with a } -\text{ve error} < 100Y \cos \theta \frac{X}{1 - X} \text{ per cent}$$

For the Extraordinary ray

$$\frac{X}{1 - n^2} = 1 - \frac{Y \cos \theta}{1 + q} - \frac{Y^2 \sin^2 \theta}{1 - X} \tag{5}$$

with the same (negative) error as (3), and

$$\frac{X}{1 - n^2} = 1 - (1 + q)Y \cos \theta - \varepsilon \text{ where } \varepsilon < \frac{Y \sin^4 \theta}{8 \cos \theta} \tag{6}$$

For several values of X, equations (3) and (4) are shown in a polar diagram as $n(\theta)$ in Fig. 1 using the value $Y = 0{\cdot}24$ for which the exact values of n (shown as solid lines) have been calculated by Forsgren (1951). The "longitudinal approximation" of equation (4) (broken lines) is seen to be sufficiently accurate for most purposes for values of X up to at least 0·6 for any direction of magnetic field. It is thus applicable to normal oblique incidence transmission with angles of incidence greater than 40°. Since the error is roughly proportional to $Y^2/(1 - X)$ we may expect this simple form to hold quite well for values of X up to 0·9 at frequencies of 10 Mc/s and above. The dashed lines represent equation (3) and are reasonably accurate for all values of X and θ. Both equations are exact in the limiting cases of $\theta = 0$, $\pi/2$ or $X = 0$, 1 or $Y = 0$.

Argence and Mayot (1953) used an approximation

$$n^2 = \frac{(1 - X)(1 - X + Y)}{1 - X + Y + XY \cos^2 \theta}$$

to calculate the shape of the virtual height curve for a parabolic layer. This was derived by approximating to the curves of the $n(\theta)$ diagram by ellipses with the

VOL. 14 1959

same axes, that corresponding to $X = 0{\cdot}9$ is shown dotted in Fig. 1. The mean error is about twice that for equation (3) and decreases with X, becoming half that for equation (4) and several times that for equation (3) at $X = 0{\cdot}6$. In the absence of any magnetic field Fig. 1 would consist of circles with the minor axes of the present curves as diameter.

Since for a horizontal ionosphere $n \sin i$ is constant and equal to $\sin i_0$ (where i is the angle between the wave-normal and the vertical) the values of n for points along a ray path will lie on a vertical line at a horizontal distance $\sin i_0$ from the

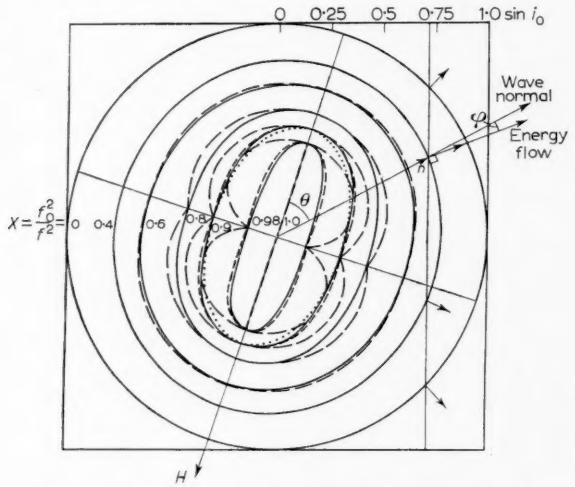


Fig. 1. Ordinary ray refractive index at $Y = 0{\cdot}24$.
——— Exact value — — — $X/1 - n^2 = 1 + Y\cos\theta$.
..... Ellipse - - - $X/1 - n^2 = 1 + Y\cos\theta/1 + q$.

centre of the appropriate $n(\theta)$ diagram (Fig. 1). This forms the basis of POEVERLEIN'S graphical method of constructing the ray paths, since the direction of propagation of the ray is perpendicular to the n-curve at a given point. The angle ϕ it makes with the wave-normal is therefore given by

$$\tan\phi = \frac{1}{n}\frac{dn}{d\theta} = \pm\frac{Y\sin\theta(1-n^2)}{2(1-X)\sqrt{(1+q^2)}} \tag{7}$$

Equation (2) then gives

$$\tan\phi = \pm\frac{Y\sin\theta}{2(1-X)}\frac{X(1+q)}{D}$$

where, for the ordinary ray,

$$D = 1 + q + q^2 + Y\cos\theta - \xi(1 - qY\cos\theta) \fallingdotseq (1+q)^2$$

giving

$$-\tan\phi_0 = \frac{XY\sin\theta}{2(1-X) + Y\sin\theta\tan\theta} \tag{8}$$

with an error between -25 per cent and $+100Y\cos\theta(1-\frac{1}{2}\tan^2\theta)$ per cent. For $X \leqslant \frac{1}{2}$ (oblique incidence) q is small unless $\cos\theta$ is small, and

$$\frac{D}{1+q} \fallingdotseq 1 + q + \frac{Y\cos\theta - q}{1+q}$$

giving

$$-\tan\phi_0 = \frac{XY\sin\theta}{2(1-X) + Y\sec\theta} \tag{9}$$

with an error between $-100q/(1+q)^2$ per cent and $+50Y\cos\theta(1-\tan^2\theta)$ per cent.

The deviations due to the magnetic field are found by evaluating $\int \tan\phi\, ds$ along the ray path. The maximum positive errors given occur at the base of the ionosphere where $\tan\phi$ is zero, and disappear when $\tan\phi$ becomes appreciable. The value q is small over most of a ray path, becoming infinite at reflection (when the expressions for $\tan\phi$ are exact) for the case of vertical incidence. Thus the error will approach -25 per cent only over a small range as q passes through unity, and the mean integrated error is generally less than half this.

Differentiation of equation (4) gives

$$\tan\phi_0 = \frac{-XY\sin\theta}{2(1-X)} \tag{10}$$

The error is less than 20 per cent for $\theta < 70°$, $X < 1 - 2Y$.

For the extraordinary ray reflection occurs at $X = 1 - Y$, and q has a maximum value of $\frac{1}{2}\sin\theta\tan\theta$. Then

$$D \fallingdotseq 1 + q - Y\cos\theta(1+q)^2$$

giving

$$\tan\phi_x = \frac{XY\sin\theta}{2(1-X)(1-Y\cos\theta) - Y^2\sin^2\theta} \tag{11}$$

with an error +ve for $q < \frac{1}{3}Y$ of less than $+100q^2/(1+q)$ per cent for $Y \leqslant \frac{1}{2}$. The peak error, attained at reflection, is about 0·4 per cent at $\theta = 20°$, 2 per cent at 30°, 5 per cent at 40° and 30 per cent at $\theta = 60°$.

3. Ray Paths at Vertical Incidence

Writing

$$a = q(1-X) = \tfrac{1}{2}Y\sin\theta\tan\theta$$

and

$$a' = \frac{-aY\cos\theta}{1 - Y\cos\theta}$$

equations (8) and (11) become:

$$-\tan\phi_0 = \frac{aX\cot\theta}{1 - X + a}$$

VOL. 14 1959

and

$$-\tan \phi_x = \frac{a' X \frac{1}{Y} \operatorname{cosec} \theta}{1 - X + a'}$$

At vertical incidence θ is constant and equal to $\pi/2 - I$ where I is the magnetic dip angle.

Measuring h vertically and S in the magnetic meridian, the ray path is given by $S = \int_0^h \tan \phi \, dh$, the origin being taken at the point where the ray enters the ionosphere. For a linear layer and a unit height of reflection, we have $h = X$ giving for the Ordinary ray:

$$S = a \tan I \left\{ (1 + a) \ln \left(\frac{1 + a}{1 + a - h} \right) - h \right\} \tag{12}$$

The Extraordinary ray path is obtained by replacing the constants a and $\tan I$ by a' and $1/(Y \cos I)$ respectively.

For a parabolic layer, $X = (1/z^2)(1 - h^2)$ where h is measured downwards from the centre of the layer, the base being at $h = 1$, and z is the ratio of the wave frequency to the critical frequency of the layer. The ray path is then given by:

$$S = a \tan I \left(H + \frac{1}{H} \right) \left(\cot^{-1} H - \cot^{-1} \frac{H}{h} \right) - a \tan I (1 - h) \tag{13}$$

or

$$S = a \tan I \left(\frac{1}{K^2} - 1 \right) \left\{ h \phi \left(\frac{K}{h} \right) - \phi(K) \right\} - a \tan I (1 - h)$$

where $H^2 = -K^2 = z^2(1 + a) - 1$ and $\phi(K) = K \tanh^{-1} K$.

The path extends from $h = 1$ to reflection at $h = \sqrt{(1 - z^2)}$ (Ordinary ray) or $\sqrt{\{1 - z^2(1 - y)\}}$ (Extraordinary ray). For a layer of semithickness Δ the actual deviation is $S\Delta$ and is about twice as large (for $z \sim 1$) as for a linear layer giving the same height of reflection. The Ordinary ray path is perpendicular to the direction of the magnetic field at reflection when $\tan \phi$ has a maximum value of $\tan I$.

In Fig. 2 some ray paths derived graphically by FORSGREN (1951) are compared with the corresponding paths calculated from equation (13). The deviations always agree to within 10 per cent. The extraordinary ray curves show no significant differences, since with $\theta = 13°$ the maximum error in equations (11) and (13) is only 0·06 per cent.

4. The Horizontal Displacement of the Point of Reflection

Sweep frequency ionospheric records will not represent the variations in the ionosphere in a purely vertical direction, due to the large, varying, horizontal displacements of the reflection point near a critical frequency. For the Ordinary ray the present analysis permits a direct presentation of the magnitude and variation of this deviation, since equation (8) gives $\tan \phi \cot I$ as a function of X and of a only. Thus if S is the deviation of the point of reflection, equation (12) gives for a linear layer

$$S' = \frac{S}{h_m} \cot I = a(1 + a) \ln \left(1 + \frac{1}{a} \right) - a \tag{14}$$

VOL
14
195

S' is the ratio of the actual deviation to that which would be obtained if the ray was perpendicular to the magnetic field throughout its path.

For a parabolic layer it is convenient to use $S' = S \cot I/\Delta$, equation (13) giving for $z = 1$

$$S' = \sqrt{a}(1 + a) \cot^{-1} \sqrt{a} - a \tag{15}$$

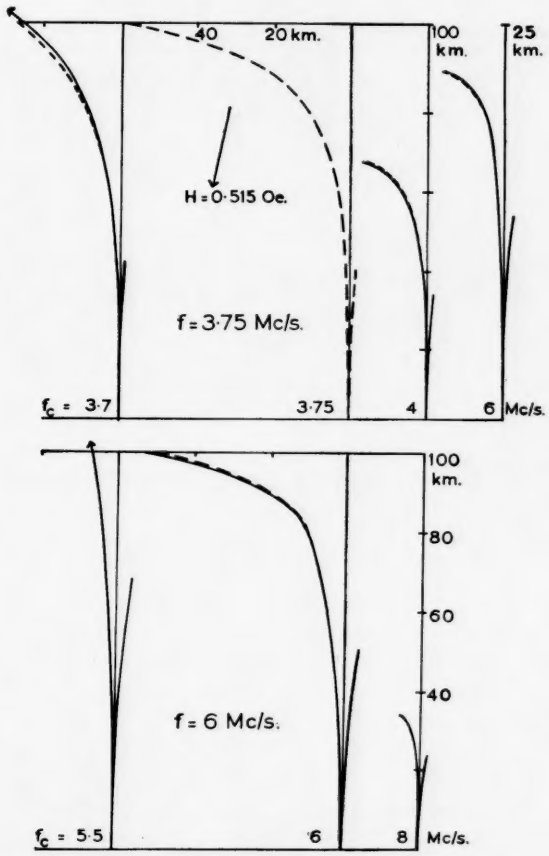


Fig. 2. Ray paths at vertical incidence for a parabolic layer with a semithickness of 100 km and critical frequency f_c.
Dip angle $I = 77°13'$; $f_H = 1{\cdot}44$ Mc/s.
————Ray paths constructed graphically by FORSGREN.
– – – – Ray paths given by equation (13) (shown only where they differ appreciably from the graphical values).

S' is plotted in Fig. 3 as a function of a for values of z equal to 0·8, 0·95, 0·99, 1·0 and 1·1. These curves determine the deviation of the corresponding rays for any frequency (or value of Y) and any strength and direction of magnetic field. The values are exact at the limits (longitudinal and transverse magnetic fields) and about 10 per cent too small elsewhere for $Y \sin I$ less than 1. The total horizontal

displacement of a transmitted ray ($z > 1$) will be twice that shown, which refers to the centre of the layer.

For a given magnetic dip, the curves show directly how the deviation varies with Y, i.e. with frequency. The variation is quite slow, a change of about ten times in Y being required to double the deviation. The deviation decreases fairly uniformly as f falls below f_c, being reduced to about three-quarters, half and quarter

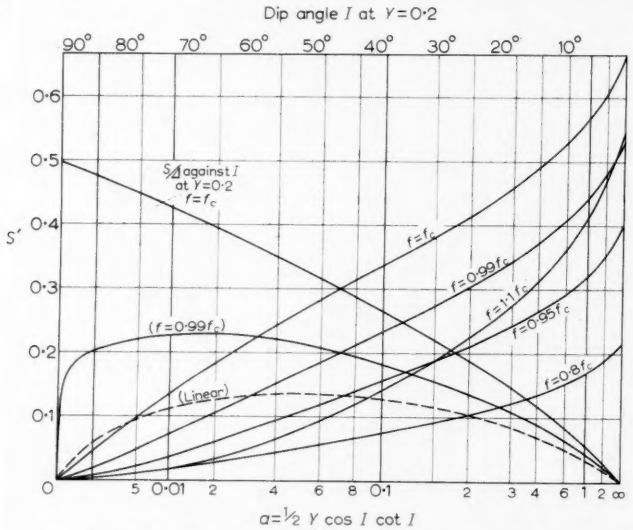


Fig. 3. Horizontal displacement of the reflection point (multiplied by cot I) for vertical incidence on a parabolic layer of unit semithickness.

of the critical value at $f/f_c = 0{\cdot}99$, $0{\cdot}95$ and $0{\cdot}8$ respectively. When $f < 0{\cdot}8f_c$ the deviation is greater than that for a linear layer (with the same height of reflection) by less than 20 per cent. When I is large the deviation mostly occurs in a narrow region near $X = 1$, and so falls off very rapidly for f greater than f_c.

For a given value of Y the curves of Fig. 3 show the variation of the displacement with magnetic dip; the values of I when $Y = 0{\cdot}2$ are shown at the top of the figure. $S/\Delta(= S' \tan I)$ is plotted against I for this case at $f = f_c$ and $f = 0{\cdot}99f_c$. The curve for a linear layer is also shown (dashed) and gives the lower limit of the ratio of the deviation to tan I times the penetration as f/f_c is reduced. It is seen that the deviation is sustained up to quite large dip angles. The limiting case $f = f_c$ (requiring an infinite virtual height and not realizable in practice) actually tends to a maximum value of $(\pi/2)\sqrt{(Y/2)}$ at the poles, where it is discontinuous. A "vertical" ray at the poles will generally make an angle of a few degrees with the magnetic field due to tilting of the reflecting layer. So for a frequency $f \fallingdotseq f_c \fallingdotseq 5f_H$ reflected from a layer of semithickness 100 km the reflection point might be expected to move on a circle of about 30 km radius.

For $f > f_H$ the curves of Fig. 3 represent approximately the lower limits of the deviations for a given value of a. The exact form of S' for the Ordinary ray is

$$S' = \int_0^1 \frac{aX}{\sqrt{\{(1-X)^2 + a^2\}}} \cdot \frac{(1-X)\,dh}{1 - X + Y \sin I[\sqrt{\{(1-X)^2 + a^2\}} - a]} \qquad (16)$$

VOL
14
195

where $1 - X = h$ for a linear layer and h^2 for a parabolic layer (at $f = f_c$). Thus the upper limit of the displacements at a given value of a are obtained at $Y \sin I = 0$, giving for a linear layer $S'_m = a\{\sinh^{-1} 1/a + a - \sqrt{(1+a^2)}\}$. These limits, and the exact values of S' (evaluated numerically) for $Y \sin I = \frac{1}{4}, \frac{1}{2}, 1$ and 2, are shown in Fig. 4. Equations (14) and (15) are shown dashed for comparison. The corrections required are seen to be of the expected order; about 10 per cent at medium frequencies. They depend only slightly on the shape of the layer,

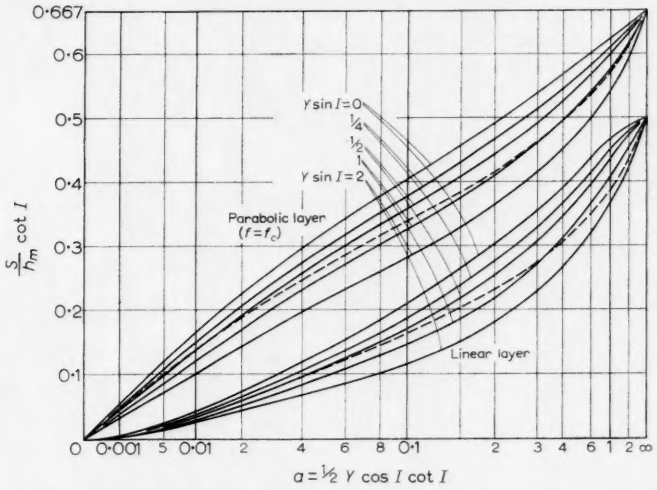


Fig. 4. Exact values of the lateral displacements for parabolic and linear layers. The dashed lines are the approximate values from equations (14) and (15).

and are readily estimated for the intermediate cases of reflection from a parabolic layer with $f < f_c$. Similar corrections will apply to the deviations obtained by integrating equation (10) for any other layer shape.

Table 1 gives the absolute corrections to equations (14) and (15) in units of the third decimal place, and may be used when an accuracy of a few per cent is insufficient. The maximum error in this table is 0·7 units.

At most observatories the magnetic dip angle is greater than 50° and equations (12) and (13) will give the deviations of the Extraordinary ray to within about 1 per cent. For equatorial regions a much closer upper limit is obtained by writing $Y \cos\theta\{q + \sqrt{(1 + q^2)}\} \fallingdotseq Y$ in equation (1). Equation (5) then gives:

$$S_m = \frac{\Delta Y \cos I}{1 - Y} \int_0^1 \frac{X\, dh}{\sqrt{\{(1 - X)^2 + a^2\}}}$$

$$= \frac{\Delta Y \cos I}{(1 - Y)^2} \left\{ \sinh^{-1} \frac{1}{a} - \sinh^{-1} \frac{Y}{a} + \sqrt{(Y^2 + a^2)} - \sqrt{(1 + a^2)} \right\} \qquad (17)$$

for a linear layer. A crude lower limit for S is given by:

$$\frac{1 - Y}{1 - Y^2} S_m$$

since $Y \cos\theta\{q + \sqrt{(1 + q^2)}\} \geqslant Y^2$.

Since for small dip angles the Extraordinary ray deviation is small and not greatly dependent on layer shape these relations should provide sufficient accuracy when equation (11) is not suitable.

A third, so-called Z-trace is occasionally observed on ionospheric records. ELLIS (1957) has shown how measurement of the penetration frequencies of the O-, X- and Z-rays can yield information on the strength of the magnetic field in

Table 1. Corrections to calculated values of S^1 ($\times 10^3$)

$Y \sin I =$	Parabolic layer ($f = f_c$)					Linear layer				
	0	$\frac{1}{4}$	$\frac{1}{2}$	1	2	0	$\frac{1}{4}$	$\frac{1}{2}$	1	2
$a = 0$	0	0	0	0	0	0	0	0	0	0
0·01	+26	+11	−1	−19	−40	+6	+1	−3	−8	−13
0·0316	+46	+25	+9	−15	−47	+18	+7	−1	−12	−23
0·1	+63	+37	+17	−14	−59	+46	+23	+6	−18	−47
0·316	+74	+51	+29	−2	−49	+82	+56	+35	+2	−44
1	+54	+40	+28	+6	−30	+81	+65	+50	+24	−18
∞	0	0	0	0	0	0	0	0	0	0

the ionosphere and on the horizontal gradient of ionization. The calculations require a knowledge of the lateral deviations of the reflection points for the three rays at a range of frequencies. Since such "triple splitting" is only observed when the magnetic dip angle is greater than 65°, equation (13) gives the deviation of the X-ray to within 0·5 per cent. The Z-ray is thought to consist of a slightly oblique Ordinary ray, becoming parallel to the magnetic field at the normal reflection level $X = 1$. It then continues in the Extraordinary mode and is reflected at $X = 1 + Y$. The refractive index is given by $n^2 \simeq 1 - X/(1 + Y)$ over the complete path, since the propagation is exactly longitudinal at $X = 1$. For the same reason the angle between the ray direction and the wave-normal is always small and may be neglected. The lateral deviation of the reflection point is then given by

$$S = h_0 \tan i_0 + \Delta \sin i_0 \operatorname{sech}^{-1} \sin i_0 \qquad (18)$$

when h_0 is the height of the base of the parabolic layer of semithickness Δ, and i_0 is the angle of incidence given by

$$\sin i_0 = \sqrt{\left(\frac{Y}{1 + Y}\right)} \cos I$$

This result is for a ray of frequency $f_z = f_c/\sqrt{(1 + Y)}$. The true critical ray with an infinite virtual height will have a frequency $f_z \sec i_0$. This will also have an infinite lateral deviation, since there is no cusp at reflection. However the back-scattering which returns the ray to the transmitter will take place somewhere below the layer maximum, and some compromise deviation such as that given by equation (18) must be used. This deviation may be too small, since the ray considered only penetrates a distance of (typically) 0·85 Δ; it does however reproduce the deviations given by ELLIS (at $I = 72°$) to within 3 per cent, the accuracy to which his curves may be read.

5. Deviations at Oblique Incidence

In the presence of a magnetic field, Snell's law still governs the direction of the wave-normal. Below the ionosphere this coincides with the ray direction, and so the sole effect of the magnetic field on a ray reflected from the ionosphere is to displace the point of emergence. $Y \sin\theta$ may be regarded as the component perpendicular to the wave-normal of a vector $\boldsymbol{Y}$ along the direction of the earth's magnetic field. The direction of the wave-normal remains in the plane of incidence so the deviation ϕ_2 of the ray direction from this plane is, by equation (7)

$$\tan\phi_2 = \pm\frac{Y_2}{2(1-X)}\frac{1-n^2}{\sqrt{(1+q^2)}}$$

where Y_2 is the horizontal component of $\boldsymbol{Y}$ perpendicular to the direction of propagation. If i is the angle between the wave-normal and the vertical, the ray is deviated out of the plane of incidence a distance $ds = \tan\phi_2/\cos i\, dh$ in a vertical distance dh. Since the deviations due to the magnetic field are small, the total displacement may be evaluated by integrating along the "wave-path" defined by $n\sin i = \sin i_0$ instead of along the actual ray path. The lateral deviation at any point is then given by

$$S = \int_1^{n^2}\frac{\tan\phi_2}{\cos i}\frac{dh}{dn^2}\,dn^2 = \pm A\,Y_2\int_1^{n^2}\frac{dh}{dn^2}\frac{1-n^2}{\sqrt{(n^2-\sin^2 i_0)}}\,dn^2$$

where A is a mean value of $n/\{2(1+X)\sqrt{(1+q^2)}\}$ over the range of integration. Writing $n^2 - \sin^2 i_0 = u^2\cos^2 i_0$ gives

$$S = \pm 2A\,Y_2\cos^3 i_0\int_1^u\frac{dh}{dn^2}(1-u^2)\,du \tag{19}$$

Equation (1) may be written

$$\frac{X}{1-n^2} = 1 + Y_1 \tag{20}$$

where Y_1 [equal to $Y\cos\theta\{-q \pm \sqrt{(1+q^2)}\}$] is small and slowly varying over a ray path. Thus for a linear layer dh/dn^2 is approximately constant and equal to $-(dh/dX)(1+Y_1)$, giving the lateral deviation at the point of reflection ($n = \sin i_0$, $u = 0$) as

$$S = \pm\frac{4}{3}A\frac{dh}{dx}(1+Y_1)Y_2\cos^3 i_0 \tag{21}$$

For a parabolic layer

$$h = \Delta\sqrt{\left(1-\frac{f^2}{f_c^2}X\right)}$$

where h is measured downwards from the centre of the layer, and writing

$$z = \frac{f\cos i_0}{f_c}\sqrt{(1+Y_1)}$$

we get

$$\frac{dh}{dn^2} = \frac{\Delta^2}{2h}\frac{z^2}{\cos^2 i_0}$$

VOL. 14 1959

Integrating to the top of the wave-path ($u = 0$) for $z \leqslant 1$, or to the centre of the layer ($u = \sqrt{(1 - 1/z^2)}$) for $z \geqslant 1$, equation (19) gives the lateral deviation over half the ray path as

$$S = \pm A \Delta B(z) Y_2 \cos i_0 \tag{22}$$

where

$$4B(z) = \left(z + \frac{1}{z}\right) \ln \left|\frac{1 + z}{1 - z}\right| - 2$$

The accuracy of equations (21) and (22) is limited by the accuracy to which the mean value A of $n/\{2(1 - X)\sqrt{(1 + q^2)}\}$ can be estimated. The changes in this function are small, lying between those of $\sqrt{(1 - X)}$ and $1/\sqrt{(1 - X)}$. For cos $i_0 = 0$, 0·2, 0·4 and 0·5 the minimum values sin i_0 of n are 1, 0·980, 0·9165 and 0·80. Since tan ϕ is zero at $n = 1$, and smoothly approaches a maximum at $n = \sin i_0$ (where the integrand of equation (19) as a function of n is infinite), mean values

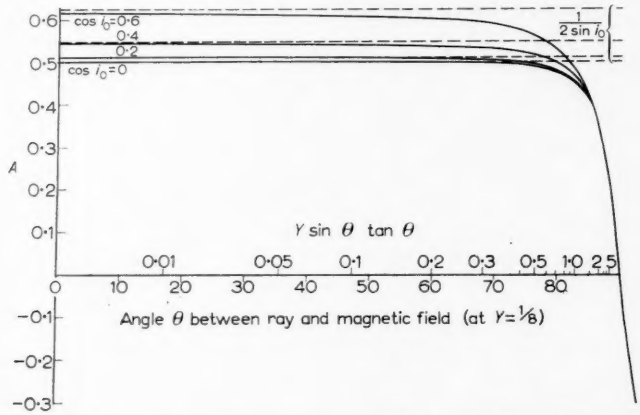


Fig. 5. Lateral deviation at oblique incidence. The actual deviation is proportional to $A \times$ (transverse component of Y).

over the ray path of 1, 0·985, 0·94 and 0·85 for n should not be in error by more than about one-third of the difference between successive values. Using

$$X = (1 - n^2)(1 + Y \cos \theta) \quad \text{and} \quad q = \frac{Y \sin \theta \tan \theta}{2(1 - X)}$$

we then get the values of A plotted in Fig. 5 as a function of θ for $Y = \frac{1}{8}$, [$f \fallingdotseq 10$ Mc/s]. The maximum error is seen to be about 2 per cent.

θ will in practice vary over the range of integration. However for $\theta < 80°$ the variation of A is small and the value of θ applicable to the horizontal direction of propagation is used. If θ becomes greater than 85° the path is best considered in two halves and mean values of θ estimated for each section. The possible percentage error in A for the section in which θ approaches 90° will be large, but the deviation is less and the overall error can still be kept small. The variation of A with θ depends almost entirely on $Y \sin \theta \tan \theta [= 2q(1 - X)]$ since $n^2 \fallingdotseq 1 - X$. Thus for any other value of Y, A may be found from the same figure using the scale of $Y \sin \theta \tan \theta$. For $\theta < 75°$, A is seen to be nearly equal to $1/(2 \sin i_0)$

(Fig. 5). The lateral deviation $2S$ over a complete path is then given by

$$2S = \Delta \, . \, B(z) Y_2 \cot i_0$$

for a parabolic layer and

$$2S = \frac{4}{3}(1 + Y_1)\frac{dh}{dx} Y_2 \cos^2 i_0 \cot i_0 = \frac{4}{3} h_m Y_2 \cot i_0 \qquad (23)$$

for a linear layer, where h_m is the height of reflection.

The values plotted in Fig. 5 are for the Ordinary ray. For the Extraordinary ray the values of A differ by less than the change in $\sqrt{(1 - X)}$, or by a factor less than $1 + Y \cos \theta$. However at a given frequency the displacement of the Extraordinary ray is less due to the smaller penetration. The Ordinary ray deviations are always such as to increase the acute angle between the directions of the ray and the magnetic field, while the deviation of the Extraordinary ray decreases this angle.

CHATTERJEE (1952) has graphically estimated the displacements over two paths by applying BOOKER's method (1949) to the special cases of equatorial and (magnetically) east–west transmission. For the equatorial path Calcutta–Bandoeng he took $f_H = 1{\cdot}0$ Mc/s, $f = 9{\cdot}53$ Mc/s and $i_0 = 70° \ 21'$. The angle between the magnetic meridian and the plane of incidence was 35° and a linear approximation, with a gradient equal to that measured at the reflection point was used. This gradient was not given, however assuming $(f \cos i_0)/f_c = 0{\cdot}9$ (i.e. a night-time critical frequency of 3·5 Mc/s) a parabolic layer of semithickness 100 km gives $dh/dX(1 + Y_1)$ equal to 940 and a total deviation of 2·9 km from equation (23). CHATTERJEE gives his result as about 3 km. In the second case (Calcutta–Bombay) $I = 30°$ and $i_0 = 65° \ 48'$. This illustrates the difficult case (of less practical importance) in the present analysis when $\theta = 90°$ at the reflection point and there is no net deviation. At the base of the layer $Y \sin \theta \tan \theta = 0{\cdot}45$. For a linear layer Fig. 5 then suggests a mean value of from 0·2 to 0·3 for A, and equation (21) gives the horizontal displacement at the centre of the path as 1·6 to 2·4 km. CHATTERJEE's result was "estimated to be about 2 km."

These figures are typical, and only for frequencies below 5 Mc/s will the total displacement over a ray path normally exceed 10 km. The largest deviations occur for an equatorial path inclined at 20–80° to the horizontal direction of the field, when the apparent bearing of a 10 Mc/s wave will be altered by about 0·1–0·3° per hop. The effect is roughly proportional to $Y_2 \cot i_0$ or $1/f^2$, since the elevation angles in normal short wave transmissions are such that $f \cos i_0$ is about the same at all frequencies.

The increased range (for the Ordinary ray) in the presence of a magnetic field is largely due to the greater penetration of the "wave-path" $n \sin i = \sin i_0$, since X is now greater than $1 - n^2$. Equation (20) gives:

$$n^2 = 1 - \frac{X}{1 + Y_1} = 1 - \frac{{f_0}^2}{f^2(1 + Y_1)}$$

and the refractive index at any point is the same as in the absence of a magnetic field if the critical frequency at that point is divided (or the effective wave-frequency multiplied) by $\sqrt{(1 + Y_1)} \fallingdotseq 1 + \frac{1}{2} Y \cos \theta$. This apparent change then

VOL. 14 1959

gives the deviation of the wave-path, caused by the magnetic field, for any form of ionosphere. For a linear layer the height of a given ionization (i.e. of f_0^2) is effectively multiplied by $1 + Y_1$, and the effect of the magnetic field is to increase the scale of the ray-diagram, and the range in the layer, by about $\pm 100\ Y \cos\theta$ per cent. For a parabolic layer the effect is equivalent to a reduction of about $\pm 50 Y \cos\theta$ per cent in the critical frequency, where θ may be taken as the (acute) angle between the horizontal direction of propagation and the magnetic field. At oblique incidence this will fairly accurately give the total increase in range, since the variation of $\cos\theta$ and the deviation of the ray from the wave-path are small. At grazing incidence the increase is exactly given by the apparent decrease $\{1 - 1/\sqrt{(1 + Y_1)}\} f_c$ in critical frequency.

POEVERLEIN (1950) gives ray paths for a linear layer with $Y = 0{\cdot}375$, $I = 65°$ and the direction of the magnetic field lying in the plane of incidence. Since $\cos\theta$ ($=\cos I$) is small the full expression (20) must be used, and the factor $1 + Y_1$ gives an increase in range (exact at grazing incidence) of 11·1 per cent. Measurement of POEVERLEIN's curves gave increases of 22 per cent at $i_0 = 18°$, 18 per cent at 30°, 15 per cent at 36°, 13 per cent at 42° and 12 per cent at 51°, the largest angle of incidence for which a ray path was given. It thus appears that the simple mean wave-path increase calculation is adequate for angles of incidence greater than about 45°, the smallest likely to be encountered in normal short-wave transmission. Since the increases due to the wave-path and due to the deviations of the ray from this path are both roughly proportional to Y, changes of frequency (above about 2 Mc/s) will not invalidate this result. For smaller angles of incidence a near-vertical magnetic field always increases this approximate range, a horizontal field in the direction of propagation decreases the range while a horizontal field at right angles to this direction has no effect.

REFERENCES

ARGENCE E.	1955	*Report of Conference on Ionosphere* p. 288. Physical Society, London.
ARGENCE E. and MAYOT M.	1953	*J. Geophys. Res.* **58,** 147.
BOOKER H. G.	1938	*Phil. Trans.* A **237,** 425.
BOOKER H. G.	1949	*J. Geophys. Res.* **54,** 243.
CHATTERJEE B.	1952	*Indian J. Phys.* **26,** 297.
ELLIS G. R.	1957	*J. Atmosph. Terr. Phys.* **11,** 54.
FORSGREN S. K. H.	1951	*Rep. Res. Lab. Electron. Chalmers Univ. Tech.* No. 17.
HASELGROVE J.	1957	*Proc. Phys. Soc.* B **70,** 653.
POEVERLEIN H.	1950	*Z. angew. Phys.* **2,** 152.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 63 to 81. Pergamon Press Ltd. Printed in Northern Ireland

Observations of abnormal VHF radio wave absorption at medium and high latitudes*

G. C. Reid† and C. Collins
Radio Physics Laboratory, Defence Research Board, Ottawa, Canada

(*Received* 19 *May* 1958)

Abstract—A study of cosmic noise absorption at a frequency of 30 Mc/s at Ottawa and Churchill has revealed the existence of two apparently distinct types of abnormal absorption event. One of these is predominantly a night-time phenomenon and is closely associated with auroral and geomagnetic disturbance. It is suggested that this absorption may be caused by an increase in electron collisional frequency at *E*-region heights rather than by a large increase in electron density at lower levels. The second type of absorption is confined to the auroral zone and is predominantly a daytime phenomenon, recurring for several days after a large solar flare. Evidence is presented to show that this absorption is due to an increase in ionization at very low levels in the ionosphere. The cosmic noise measurements are supported by evidence from a number of VHF forward-scatter circuits in Canada, and this is used to obtain information about the geographical extent and frequency of occurrence of these abnormal absorption events.

VOL.
14
1959

Introduction

It has been realized within recent years that a great deal of useful information about conditions in the lower part of the ionosphere can be obtained by the measurement of the absorption suffered by the very-high-frequency radio waves which form part of the continuous-frequency spectrum of galactic or cosmic noise. The method was first used by Mitra and Shain (1953) working at a frequency of 18·3 Mc/s and has been further developed since, notably by workers at College, Alaska (Little, 1954; Little and Leinbach, 1958), who have used frequencies of 30 and 65 Mc/s to study absorption in the auroral zone.

In principle the method is extremely simple, the basic equipment necessary being a broad-beam antenna directed towards the zenith, feeding into a preamplifier and low-noise receiver. In the absence of interfering signals, a pen recorder on the receiver output will give a continuous record of the incident cosmic noise signal within the receiver bandwidth. However, it is difficult to eliminate slight drifts in the receiver gain characteristics, which would make accurate measurement of signal strength impossible without frequent calibration. A more refined method of measuring the incident cosmic noise power, originally due to Machin *et al.* (1952), has recently been developed by Little and Leinbach (1958) and the present equipment has been based on their design.

The input to the receiver is switched at a frequency of about 100 c/s between the antenna and the output of a local noise-generating diode. If these two signals are of different amplitudes, a 100 c/s component will be present at the receiver output, having an amplitude which is proportional to the misalignment between the two inputs and being either in phase or 180° out of phase with the switching oscillator output, depending on which of the two inputs is the larger. This 100 c/s signal is amplified and detected in a phase-sensitive detector, giving a d.c. voltage

* Work carried out under Project No. PCC 48-28-01-02.
† Present address: Geophysical Institute, University of Alaska, College, Alaska.

whose amplitude and polarity are determined by the misalignment between the two inputs. This voltage is then fed back to the local noise source in such a way as to alter the noise diode output towards equality with that of the antenna. The plate current of the noise diode, which is proportional to the noise power supplied by the diode, is then recorded. By this means the effect of drift in the receiver characteristics can be largely eliminated.

Equipment of this type has been in continuous operation at a site near Ottawa since March 1957, and at Fort Churchill, Manitoba, close to the auroral zone maximum, since June 1957. Both equipments operate on a frequency close to 30 Mc/s, using dipole arrays having a 60° beam-width between half-power points, centred on the zenith.

In order to determine the absorption present at any time it is necessary first to find the undisturbed cosmic noise level, i.e. the signal which would have been recorded at that time in the absence of any absorption. This can be done by examining the records for several days on either side of the period considered, taking into account the shift of sidereal time with respect to solar time of about 4 min/day. Tabulated values of absorption at a frequency of 2 Mc/s, as recorded by ionospheric sounders at Ottawa and Churchill, have also proved extremely useful in this connexion. Since the quantity recorded is proportional to the noise power supplied by the antenna, the absorption (dB) is given simply by

$$A = 10 \log (R_0/R)$$

where R is the signal actually recorded and R_0 is the signal normally recorded at the same sidereal time. It is estimated that the accuracy of measurement of absorption is of the order of $\pm 0{\cdot}1$ dB.

It should be pointed out that the normal daytime D-region absorption, showing the familiar $\cos \chi$ diurnal variation, is a fairly small effect at frequencies of the order of 30 Mc/s. The work of Appleton and Piggott (1954) has shown that at a frequency of 4 Mc/s this absorption should have a midsummer noon value of about 20 dB at middle latitudes. Since the amount of non-deviative absorption is inversely proportional to the square of the frequency, this would correspond to about 0·35 dB at 30 Mc/s. Even if this regular absorption were to increase by a factor of three near sunspot maximum, the absorption at 30 Mc/s would only reach a value of about 1 dB at midsummer noon. Since this absorption is already present in the reference level, as found from the comparison records, its effect is ignored in the present work. We are concerned here only with departures from this regular absorption.

Abnormal Absorption Phenomena

The absorption phenomena observed can be classified into three main types:

(1) Absorption occurring at the same time as major solar flares. This type of absorption has been recognized for many years, since it causes sudden fade-outs on short-wave radio links at all points on the sunlit hemisphere of the earth. It has long been known as "sudden ionospheric disturbance" (SID) from the characteristically rapid onset of the effect. Occurrence of this type of absorption is easily recognizable, and its characteristic appearance on cosmic noise recordings has been

VOL
14
195

fully described by SHAIN and MITRA (1954). It is, of course, confined to daylight hours, being caused by enhanced short wavelength solar radiation associated with the flare, and its duration is usually less than 1 hr after the start of the visible flare. An example of this effect is shown in Fig. 1(a), the absorption in this case being preceded by an outburst of solar radio noise. SID phenomena are seen on the recordings at both Ottawa and Churchill, the magnitude of the absorption usually being less at the latter station due to the increased solar zenith angle.

(2) The second type of abnormal absorption phenomenon is for the most part a night-time effect, and its occurrence is very closely correlated with that of visible aurora and local geomagnetic disturbance. The cosmic noise recordings of this type of absorption have a characteristically irregular appearance. Although a period of absorption may last several hours, individual peaks often have durations of only a few minutes, and rates of onset and disappearance of absorption may be as high as 2–3 dB/min. This type of absorption is observed at both Ottawa and Churchill, the frequency of occurrence being greater at Churchill as would be expected from its location in a zone of greater auroral frequency. No detailed correlation studies of occurrence at the two stations have been carried out, but it is fairly evident that, while there is a definite correlation on a daily basis between occurrences at the two stations, no detailed hourly correlation between the magnitudes of the absorption exists. In general, when this type of absorption (referred to hereafter as "auroral" absorption) occurs at Churchill it does not necessarily occur at Ottawa, but occurrence at Ottawa always seems to be accompanied by occurrence at Churchill. The magnitude of the absorption can be as high as 8–10 dB, though only for short periods, and cases have occurred in which the absorption was much more intense and prolonged at Ottawa than at Churchill during a given night.

The occurrence of absorption associated with aurora at College, Alaska, has been discussed extensively by HEPPNER *et al.* (1952) and LITTLE and LEINBACH (1958). These authors find the absorption to be strongly correlated in time with the more active forms of visible aurora rather than with steady arcs and glows. This observation appears to be borne out by our own records, though no detailed study has been carried out.

An example of auroral absorption, taken from the Ottawa recordings, is shown in Fig. 1(b).

(3) The third type of abnormal absorption phenomenon differs from the others in being apparently confined to northern latitudes. The cosmic noise recordings of this type of absorption are much more uniform in appearance than the recordings of the auroral absorption described above, and magnitude of the absorption is much greater during the day than during the night. This type of absorption has appeared almost exclusively on the Churchill records, no clearly defined cases having so far appeared at Ottawa. It appears to be closely connected with great solar activity, and is very similar to the absorption described by BAILEY (1957), which occurred at northern latitudes following the great solar flare of 23 February 1956. The effects which will be described in the following section, while they are rather smaller in magnitude than those described by BAILEY, are thought to be due to the same mechanism, and they seem to suggest that such events are more frequent than has been suspected.

An example showing the increase in this type of absorption at about the time of sunrise is shown in Fig. 1(c).

On a number of occasions all three types of absorption phenomenon have been observed on the VHF scatter circuits operated by this establishment in western and north-western Canada. Some of these circuits have been in operation for a considerable period. The earlier ones were set up to study the propagation of VHF signals from ionized meteor trails, and the later ones have been used for the investigation of auroral ionization. Table 1 gives the relevant characteristics of these experimental

Table 1. VHF scatter-circuit characteristics

Path (transmitter to receiver)	Length (km)	Frequency (Mc/s)	Transmitter Power (W)	Mid-path geomagnetic co-ordinates	
				Lat.(*N*)	*Long.*(*W*)
Yellowknife, N.W.T. to Baker Lake	900	49·99	700	72	58
Yellowknife, N.W.T. to Churchill	1120	49·99	700	70	53
Yellowknife, N.W.T. to The Pas	1180	49·99	700	66	55
Yellowknife, N.W.T. to Saskatoon	1200	49·99	700	65	57
Yellowknife, N.W.T. to Sulphur Mountain	1200	49·99	700	63	62
Churchill to Winnipeg	1000	39·23	100	64	37
Suffield to Winnipeg	1000	39·21	100	59	46
Greenwood, N.S. to Ottawa	860	39·22	100	57	1

circuits. Medium-power transmitters were used to radiate an unmodulated signal which was interrupted at regular intervals for identification. Five-element yagi antennae were used at all the stations, sited at heights to illuminate the mid-path regions at approximately the 100 km level. The exception was Yellowknife where double-stacked square loop antennae were required to provide the necessary coverage. Conventional equipment was employed at the receiving stations where the amplitude of the VHF signals was recorded on moving-chart recorders with a frequency response of 2 c/s. Path lengths averaged about 1000 km.

Most of the signals propagated over these paths are forward-scattered from ionized meteor columns, but other modes of propagation are often dominant. These might be described generally as scattering modes caused by increased ionization at *E*-region heights or less. These scatter signals have been the subject of a recent extensive study by Collins and Forsyth (1958) and, while such signals do produce some contamination of the meteor signal records, they do not affect the interpretation of the records for the absorption phenomena. Fig. 2(a) shows a sample of the amplitude recording made at Winnipeg of the 39 Mc/s transmission from Churchill when propagation conditions were normal. The spiky appearance of the record is caused by the reflections from individual meteor trails. Fig. 2(b)

VOL
14
195

VOL.
14
1959

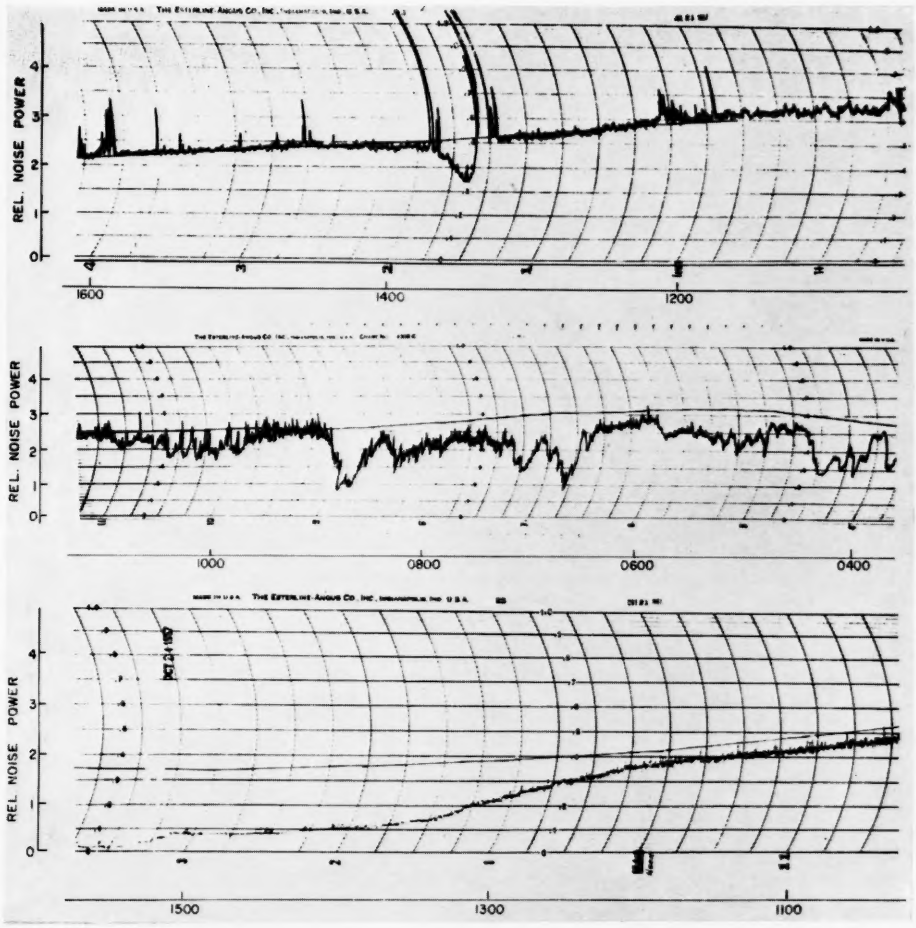
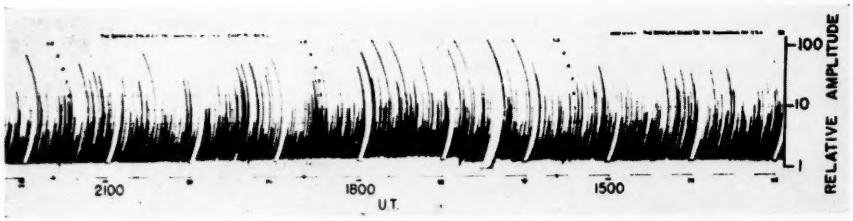
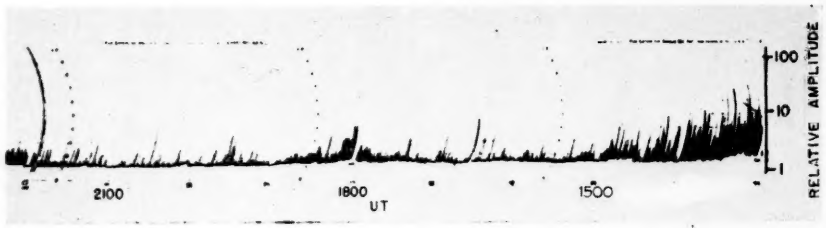


Fig. 1. Examples of abnormal absorption events. In each case the time scale is Universal Time (UT) and the thin line represents the cosmic noise signal in the absence of absorption. (a) Sudden ionospheric disturbance (SID) associated with a solar flare; (b) auroral absorption; (c) high-latitude absorption following a solar flare. Ground sunrise is at about 1300 hours.



(a)



(b)

Fig. 2. Typical VHF forward-scatter recordings. (a) Normal conditions; (b) strong daytime absorption.

shows a second recording made at Winnipeg some days later when intense absorption has caused appreciable attenuation of the meteor signals. In analysing the records no attempt was made to derive a continuous average signal level. The record for those days when absorption was believed to be present was compared to the record made on a day when propagation conditions were normal. From the general attenuation of the meteor signals the absorption was considered to be slight, if the effect was just barely apparent; moderate, if the general level had decreased about half; strong, if the signals were almost completely attenuated.

The effect of absorption at the frequencies used on these paths was first dealt with by VOGAN and CAMPBELL (1957) who showed the effect of small values of absorption on the diurnal variation of meteor signals. FORSYTH and VOGAN (1957), in discussing the frequency dependence of VHF signals propagated by auroral ionization between Greenwood and Ottawa, estimated the intensity of the effect and the height of the absorbing region, but all the events occurred during the early morning hours and almost certainly refer to what has been described earlier in this paper as auroral absorption. If the absorption takes place in an extensive layer, the amount of absorption detected on these scatter circuits will in general exceed that measured by the cosmic noise equipment since the transmission follows an oblique path and passes twice through the absorbing region. It is related to the absorption measured at vertical incidence at the same frequency by a factor equal to twice the secant of the angle of incidence. For a 1000 km path this is a factor of 8 or 9. Interpretation of the absorption events on the VHF scatter circuits was considered to be consistent with the cosmic noise results obtained at Churchill and Ottawa, and the scatter circuit data was examined with a view to obtaining some measure of the frequency of occurrence and the geographical extent of the third type of absorption phenomenon.

VOL. 14 1959

Two specific events in which this third type of absorption appeared to show some kind of systematic behaviour will now be examined. In Fig. 3 the 30 Mc/s cosmic noise absorption measured at Churchill is plotted as a function of time for the 3 day period 3–5 July 1957. The shaded strips in the centre of the diagram indicate night-time at ground level and at a height of 50 km. At this time of year in Churchill (latitude 58·8°N) the highest level in the atmosphere at which sunset occurs is about 50 km and the whole ionosphere is in perpetual sunshine. In the upper part of the diagram is a histogram of the 3 hr planetary K-indices for the same period. While these indices give a good measure of world-wide geomagnetic activity their interpretation for events at a specific location is open to some doubt, and here they have been taken as giving merely a rough measure of the geomagnetic activity at Churchill. Examination of magnetometer recordings from Churchill have since confirmed that, in this case at least, the planetary K-indices do give a reliable indication of local events.

The times shown in the diagram are all universal time (UT), which is 6 hr ahead of local standard time at Churchill. The individual points on the absorption curve are plotted at 15 min intervals.

During the early hours of 3 July magnetic activity was at a high level and small amounts of absorption were present. By 0730 this absorption had largely died out. At 0715 and 0800 (shown by arrows on the diagram) two solar flares of

importance 3 occurred, both associated with an active region in the north-west quadrant of the sun. The cosmic noise absorption started to increase at about the time of the second flare (45 min after the first), and shortly before ground sunrise the rate of increase became much greater. Superposed on this steady increase were short periods of irregular enhanced absorption, apparently associated with an increase of magnetic activity. This increased absorption is thought to be of type 2 (auroral) and its occurrence in daytime is unusual.

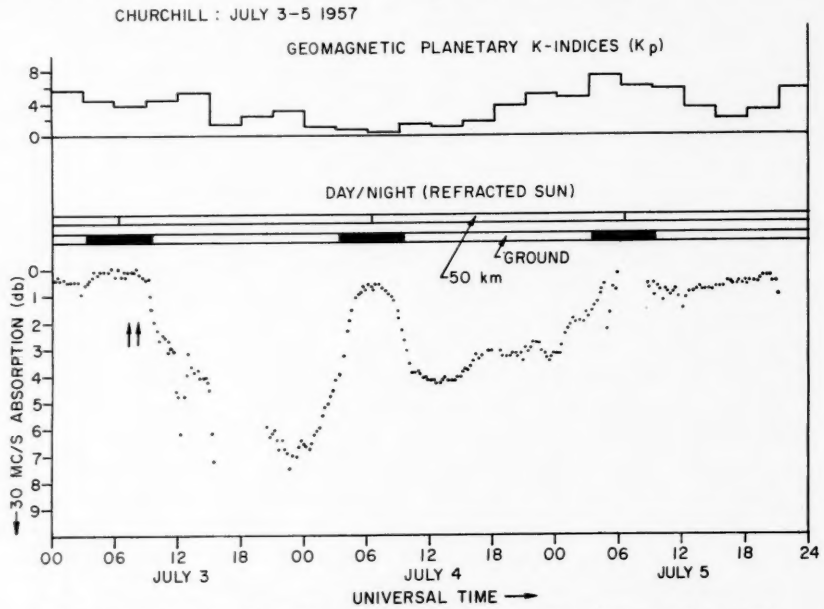


Fig. 3. Cosmic noise absorption (30 Mc/s) observed at Churchill, 3–5 July 1957.

The steady absorption appeared to reach a maximum value of 7 dB at about 2230 (1630 LT) then started to decrease rapidly, reaching one-tenth of its peak value by about local midnight. This rapid decrease seems rather difficult to understand in view of the fact that sunset did not occur at atmospheric levels which are generally thought to be responsible for VHF absorption. During the latter part of the day geomagnetic activity was at a low level.

The absorption remained at about 0·7 dB for approximately 1 hr, then started to increase again. This time the maximum value (4·5 dB) was reached at about 1230 (0630 LT), and throughout the day the absorption decreased steadily, the rate of decrease being accelerated during the evening.

At about 0430 on 5 July, auroral absorption appeared, which was associated with a large increase in magnetic activity. Absorption during the daytime on 5 July was quite low, and by evening the cosmic noise signal had almost reached its undisturbed level.

Fig. 4 shows the course of events at Ottawa during 3–5 July. The planetary K-indices are again plotted at the top of the diagram, and the shaded bars indicate night-time at ground level and at heights of 50 km and 100 km respectively.

It can be seen that the absorption at Ottawa showed no regular diurnal variation like that observed at Churchill; in fact, the absorption appears to follow quite closely the trend of geomagnetic activity as shown by the planetary K-indices. It is interesting to note that during the night of 4–5 July greater absorption was experienced at Ottawa than at Churchill, coincident with fairly intense world-wide geomagnetic disturbances ($K_p = 7+$).

The effects of this phenomenon were also observed in the signal-amplitude recordings made at all five stations which monitor the VHF transmissions from

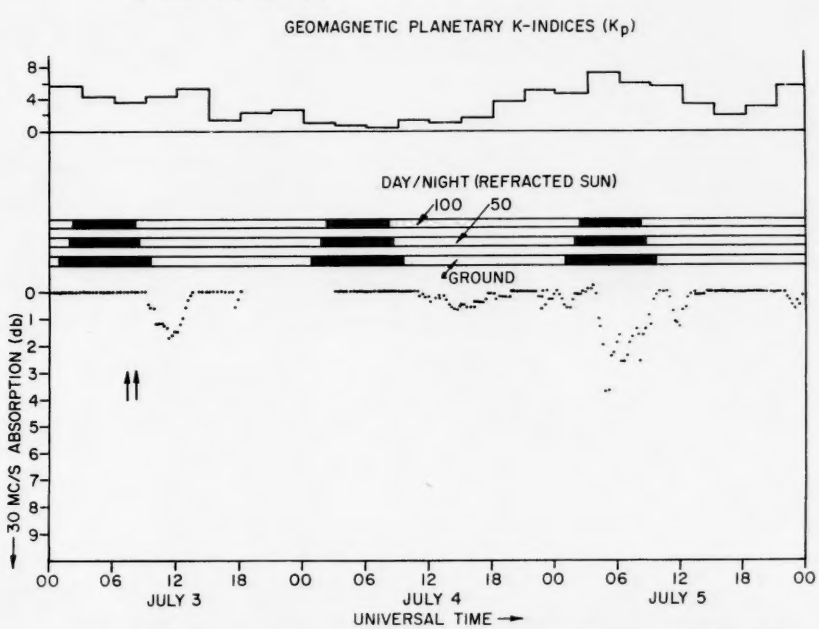


Fig. 4. Cosmic noise absorption (30 Mc/s) observed at Ottawa, 3–5 July 1957.

Yellowknife. The data from these circuits, since they showed a sequence of events in complete agreement with those deduced from the cosmic noise data, give some indication of the geographical extent of the disturbance. There was no indication on any of the records of the solar flares at 0715 and 0800. By 1045 there was slight absorption on the Baker Lake circuit, and a little later it was apparent on the Churchill, The Pas and the Saskatoon circuits. At about the same time, the low-level scatter signal, which is often propagated over these paths, was considerably enhanced. The occurrence of this scatter signal, which may be associated with the absorption phenomenon, is the "S-event" described by Collins and Forsyth (1958). With sunrise over the mid-path regions the absorption increased to about 20 dB. This is a conservative estimate but it is in reasonable agreement with the cosmic noise measurements made at Churchill. The strong absorption continued over the four paths until about sunset. On the Sulphur Mountain circuit, where the mid-path region is 700 km from the zone of maximum auroral occurrence,

VOL. 14 1959

similar effects were observed, but both the enhancement and the absorption were weaker than on the other circuits. During the night of 4 July there seemed to be slight absorption on the Baker Lake, Churchill and The Pas circuits. There was also considerable enhancement of the low-level scatter signal on the Churchill circuit and somewhat less of this signal on the other two paths. The two most southerly circuits, Saskatoon and Sulphur Mountain, appeared to have returned to normal. On 4 July, for the daylight hours before noon, slight absorption was again observed on the three paths closest to the auroral zone. Throughout the day

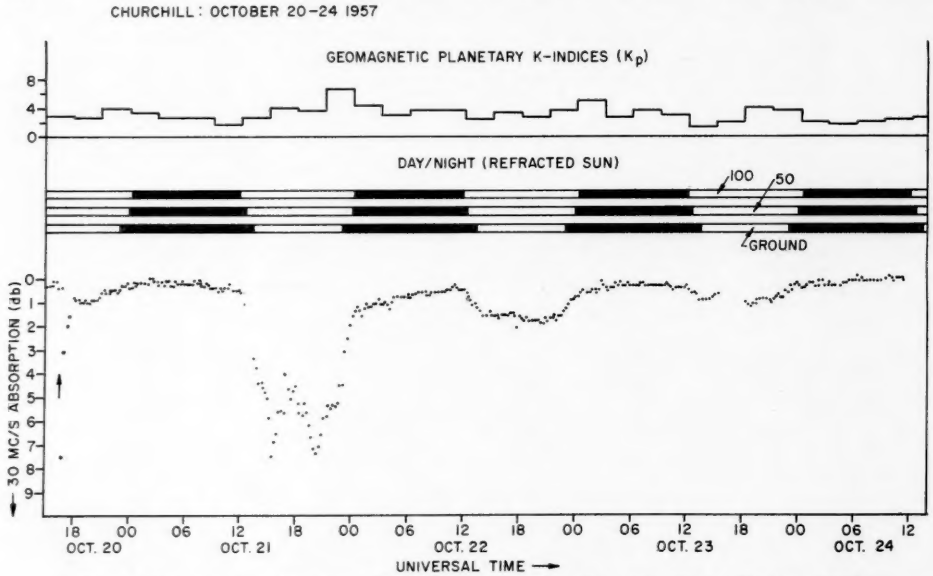


Fig. 5. Cosmic noise absorption (30 Mc/s) observed at Churchill, 20–24 October 1957.

this absorption gradually decreased while the scatter signal was at times enhanced. The Sulphur Mountain and Saskatoon circuits were normal. These effects lasted longest on the Churchill path, and there was still some absorption on the night of 4 July. At about 0415 on 5 July, a large signal enhancement occurred on all the records, of a type which Collins and Forsyth have associated with increased auroral ionization. This seemed to mark the end of the absorption phenomenon on the scatter circuits. Although the meteor trails which propagate signals over these paths may be formed over a very wide area, most of the VHF signals are scattered from trails which are formed in some broadly defined region above the middle of the path. The data from the five circuits suggests that on the first day of the disturbance the absorption was effective at least 250 km north and 700 km south of the zone of maximum auroral occurrence, and on the second day, although the effects were considerably smaller, the absorbing region still extended 250 km either side of the zone. There was no indication at any time that the phenomenon had extended as far south as the Greenwood–Ottawa path.

A second example of the same phenomenon is shown in Fig. 5. Here, absorption values at Churchill are plotted at 15 min intervals for the period 20–24 October

1957. Once again planetary K-indices and periods of day and night at ground level and at heights of 50 km and 100 km are plotted on the same diagram.

Shortly after 1630 UT on October 20th, a solar flare of importance 3^+ occurred in the south-west quadrant of the sun. The arrow indicates the start of the visual flare, and the absorption accompanying the flare (sudden ionospheric disturbance) is clearly seen at the left of the diagram. The sequence of events during the three succeeding days is seen to be similar to that of 3–5 July, but with some important differences in detail as described below.

The SID effect in its normal form had disappeared by about 1800 on 20 October but seemed to be followed by a period of continuous diminishing absorption of quite small magnitude, disappearing at about the time of sunset in the lower ionosphere. This effect is unusual, but may have been due to continued emission of the short wavelength radiation responsible for the SID.

The daytime absorption on 21 October, starting at about the time of 50 km sunrise, showed two maxima approximately equal in size and symmetrically placed about local noon (1800 UT). The night-time absorption on 21–22 October showed a steady decrease in intensity and the daytime absorption on 22 October was again roughly symmetrical about local noon. The low intensity daytime absorption on the 23 October also appeared to be symmetrical about noon.

In this case the level of world-wide geomagnetic disturbance was moderately high throughout the whole period but there appears to be very little detailed correlation between absorption and K_p.

The only unusual feature of the Ottawa recordings during this period was the occurrence of auroral absorption during the evening of 21 October coinciding with an active auroral display of short duration at Ottawa.

The effects of this second phenomenon on the 49 Mc/s transmission from Yellowknife were similar to those observed for the July event. There were however, a few different features and the effects were neither so intense, nor so extensive. The occurrence of the flare at 1635 on the 20th was shown by a short signal enhancement at Churchill, The Pas and Sulphur Mountain, which was probably due to an outburst of solar noise. This was followed about 10 min later by a short period of strong absorption, characteristic of the SID phenomenon. A high level of local interference masked the effect at Saskatoon but there is no apparent reason for it failing to be detected at Baker Lake. The period on 20 October following the outburst appears to have been normal on all circuits; the slight absorption measured by the cosmic noise equipment was too weak to affect any of the paths. At sunrise strong absorption was observed on all circuits. This condition lasted until approximately sunset when normal night-time conditions were restored. The strongest absorption, estimated at approximately 15 dB, was observed on the Baker Lake, Churchill and The Pas circuits, and a weak scatter signal was recorded late in the day at Baker Lake. At Saskatoon the absorption reached approximately 10 dB, while on the Sulphur Mountain circuit the effect was weaker and did not last until sunset. On 22 October slight absorption could be detected on the Churchill and on The Pas circuits, but the other paths appeared normal. The absorption measured on 23 October by the cosmic noise equipment was again too weak to be observed on any of the scatter circuits.

VOL. 14 1959

The data obtained from a year's operation of two other scatter circuits, Churchill–Winnipeg and Suffield–Winnipeg, were examined to obtain some indication of the frequency of occurrence of this type of absorption phenomenon. The circuits were operated from July 1956 to June 1957. The operating frequency was 39 Mc/s. These data also provide some further information about the extent of the absorbing region, since the mid-path region for the former circuit is about 500 km from the zone of maximum auroral occurrence, while the mid-path region of the latter is 1000 km from this zone. No attempt was made to scale average signal amplitudes. A selection was made of those daytime periods which showed an attenuation of the meteor signals of the order of 20 dB, starting around sunrise and decreasing about sunset. Examination of the Churchill–Winnipeg recordings showed that there were five 2 day periods of absorption similar to those already discussed. On five other days slightly less absorption was observed which did not recur, or which recurred for only a few hours on the following day. Most of these absorption events were preceded by considerable solar activity. Four of the 2 day events and one of the 1 day events were preceded by solar flares of importance 3. Of the remainder, another 1 day event could be associated with a solar flare of importance 2. The Suffield–Winnipeg data showed that only one of the 2 day events from the Churchill path had extended as far south as the Suffield path.

Discussion

Perhaps the most important feature emerging from a study of these unusual absorption effects is the distinction between the two types of absorption phenomena, described above respectively as type 2 (auroral) and type 3 (high-latitude) absorption. The features of these two classes may be summarized as follows:

Type 2. This appears at both Ottawa and Churchill and is strongly correlated with the occurrence of active forms of visible aurora. This type of absorption is observed more frequently at Churchill than at Ottawa and while its occurrence at Ottawa during a given period is usually accompanied by occurrence at Churchill, the reverse is not generally true. This absorption is predominantly a night-time phenomenon, though it is occasionally observed during daylight hours. The absorption record presents a characteristically irregular appearance and may show onsets and disappearances which often coincide closely in time with local magnetic fluctuations of the bay type. It may be mentioned that Wells (1947) has described a close correlation between individual night-time absorption events and magnetic bays at College, Alaska.

The association between radio-wave absorption and visible aurora, though by no means new, has recently received considerable support from the results of forward-scatter, or "bistatic" experiments (see Forsyth and Vogan 1957; Collins and Forsyth, 1958), and its existence now appears to be well established. There is still some doubt, however, as to the height at which the absorption occurs, and its degree of localization. This latter factor may be important in providing an explanation for the lack of radar reflections from aurora at high angles of elevation as suggested by Currie *et al.* (1953). These two factors will be discussed later.

Type 3. This class of absorption phenomenon has not appeared in the Ottawa records to date with appreciable intensity, but seems to be confined to high

latitudes, though its geographical extent may be quite large within these latitudes, as shown by the VHF forward-scatter evidence. It appears to set in within a few hours of the onset of a major solar flare and usually persists for 2–3 days, gradually declining during this period and exhibiting strong solar control in a similar fashion to the normal *D*-region variation. This absorption shows little. if any, detailed correlation with geomagnetic activity and is predominantly a daytime phenomenon, in contrast to type 2 absorption.

CHAPMAN and LITTLE (1957) have suggested a mechanism to explain the occurrence of abnormal VHF absorption in the auroral zone. The primary cause of the absorption is thought to be an increase in ionization in the lower ionosphere due to bombardment by electrons of solar origin and their associated Bremsstrahlung radiation. The electron bombardment is greater during the night, leading to greater geomagnetic disturbance, but at the same time free electrons will quickly become attached to neutral molecules, forming negative ions which will not produce radio wave absorption. During the day the decreased electron bombardment is more than offset by photo-detachment of electrons from negative ions by the action of long–wavelength solar radiation. The negative ion hypothesis is very attractive as regards an explanation of the strong solar control exhibited by type 3 absorption, but at the same time the great excess of night-time over daytime electron flux required to explain the night-time geomagnetic disturbances seems rather difficult to understand. The theory of electron bombardment does receive considerable support from the discovery of soft radiation at heights down to 50 km in the auroral zone as a result of rocket experiments carried out by MEREDITH *et al.* (1955). This soft radiation was attributed to Bremsstrahlung generated by electrons of energies between 10–100 keV bombarding higher regions of the ionosphere.

It is suggested here that some mechanism of this kind is the most likely explanation of the type 3 absorption, and would account for the difference between the daytime and night-time absorption as seen for instance at Churchill on 3–4 July (Fig. 3), when there is no appreciable change between day and night magnetic activity, and possibly also on 20–22 October (Fig. 5). However, it is thought that some other mechanism should be invoked to explain the absorption of type 2 in order to account satisfactorily for its different characteristics. This point of view receives some justification from the fact that occasionally type 2 absorption has been observed during daytime superimposed on a steady background of type 3 absorption. An example of this is shown in Fig. 6, and serves to demonstrate that the two types of absorption are certainly not mutually exclusive, from which we can conclude that the characteristic appearance of type 2 absorption is not due to the existence of night-time conditions in the ionosphere. A different mechanism to account for type 2 absorption will be described in the following section.

Type 2 (auroral) absorption

The absorption suffered by a radio wave in passing through an ionized medium can be characterized by the relation

$$I = I_0 \exp(-ks)$$

where I is the power received, I_0 is the power incident on the top of the region responsible for the absorption, s is the distance travelled through the region and k is the absorption index. The product ks is thus seen to be proportional to the absorption suffered by the radio wave measured in decibels.

In practice, k in the ionosphere varies from point to point and the quantity deduced from absorption measurements is $\int k\,ds$ along the path of the radio wave.

The factor k has been derived from magneto-ionic theory in terms of the

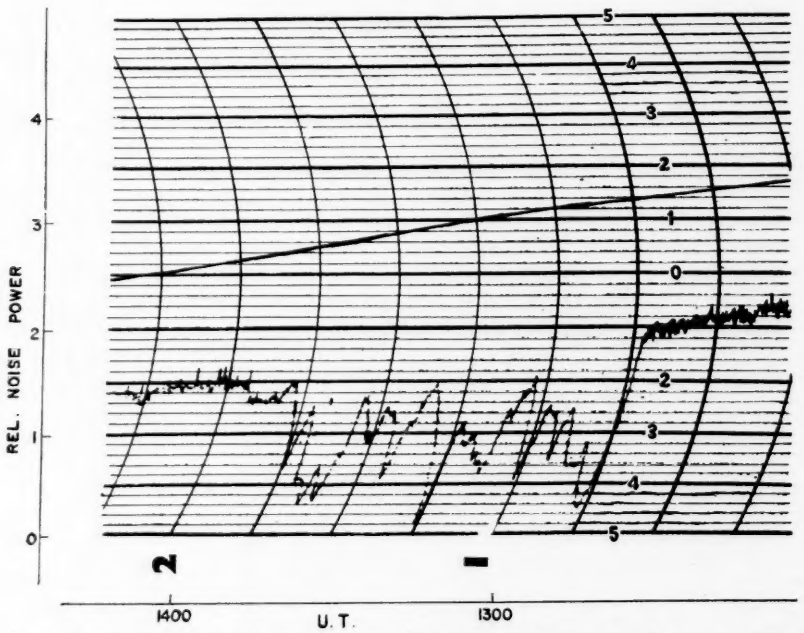


Fig. 6. Example of type 2 (auroral) absorption, superposed on type 3 absorption. The thin line represents the normal cosmic noise signal level.

refractive index of the medium, μ, the electron density, N, the electron collisional frequency, ν, the radio wave angular frequency, ω, and the angular gyrofrequency about the longitudinal component of the geomagnetic field, ω_L, according to the equation

$$k = \frac{2\pi e^2}{mc}\,\frac{1}{\mu}\,\frac{N\nu}{\nu^2 + (\omega \pm \omega_L)^2}$$

Here, e, m and c, are respectively the charge and mass of the electron and the velocity of light, and the positive and negative signs refer to the ordinary and extraordinary modes of propagation respectively.

In the present case, using a frequency of 30 Mc/s, ω_L is much smaller than ω and can be neglected; also, since the operating frequency is well above the critical frequency, $\mu \simeq 1$. Hence, approximately,

$$\int k\,ds = \frac{2\pi e^2}{mc}\int \frac{N\nu\,ds}{\nu^2 + \omega^2}$$

At heights in the ionosphere greater than about 65 km $\nu^2 \ll \omega^2$, and the

absorption (dB) becomes proportional to $\int N\nu\, ds$, i.e. to the product of electron density and electron-collisional frequency. At a frequency of 30 Mc/s the measured absorption in dB is given by

$$D = 0{\cdot}13 \times 10^{-16} \int N\nu\, ds \qquad (1)$$

where N is measured in cm^{-3}, ν in sec^{-1} and s in cm.

It is evident that abnormal absorption can be produced either by an increase in electron density in a region of fairly high collisional frequency, or by an increase in the electron collisional frequency itself. The first alternative is the one which is normally considered, but objections can be raised to this as an explanation of the irregular auroral absorption. The generally accepted time-delay between the occurrence of a solar flare and the appearance of auroral and magnetic effects on the earth is of the order of 1–2 days. Unless some very strong local acceleration mechanism is postulated it is difficult to see how particles of these velocities can penetrate much lower than about 120 km. In actual fact most visible auroral forms have a lower border at about 100 km, so that some form of local acceleration must take place, but even allowing for this it is difficult to see how these auroral particles can penetrate into the lowest ionosphere where a reasonable increase in ionization density would account for the large amounts of absorption observed. It would be very desirable if the ionization responsible for the absorption could be located at the same level as the visible aurora itself, i.e. at a height of the order of 100 km, but it is not immediately evident how this can come about. If the absorption is taken as being due to a layer 10 km thick at the 100 km level, and if the value of collisional frequency at 100 km is taken to be that given by Nicolet (1953), then the electron density required to give 5 dB absorption at 30 Mc/s is about $2{\cdot}5 \times 10^6$ electrons per cm^3. This figure exceeds by a factor of 250 the normal midnight value of E-region electron density, and by a factor of five the maximum electron density found by vertical incidence sounding during auroral displays at College as reported by Chapman and Little (1957). (It should be pointed out in this connexion that neither the vertical incidence sounding nor the absorption measurements would reveal the existence of localized patches of much higher electron density, and the figures quoted above all refer to average values over a large area of the ionosphere.) If the possibility of penetration to lower depths is accepted however, much smaller electron densities are tolerable, due to the higher collisional frequency. Chapman and Little (1957) have constructed a model ionization density distribution which will explain an absorption of 2·5 dB at 30 Mc/s on this hypothesis.

Another process which must be considered, however, is an increase in collisional frequency without any great increase in electron density. This can occur as a result of the application of an electric field to an ionized gas, as has been pointed out by Bailey (1957). Experimental work on this effect has been carried out by Crompton, Huxley and Sutton (1953), and the possible application of their results to the disturbed ionosphere is worth consideration. When an electric field is applied across an ionized gas the electron temperature increases, leading to more rapid electron agitation and greater collisional frequency with the neutral molecules. The above authors measured k_T, the ratio of electron temperature to neutral

molecule temperature, as a function of X/p, the ratio of applied electric field to gas pressure, and by adopting experimental values of the electron drift velocity found by Nielsen and Bradbury (1937) were able to calculate ν/p, the ratio of electron collisional frequency to gas pressure, as a function of X/p. Their experiments were performed in air at 15°C (288°K), and while the temperature dependence of the effect is uncertain it is thought that their results will give the correct order of magnitude for effects at a temperature of 213°K as at the 100 km level in the ionosphere. The dependence of ν/p on X/p is shown in Fig. 7; here p is measured in mm Hg, ν in $\sec^{-1}$ and X in V/cm.

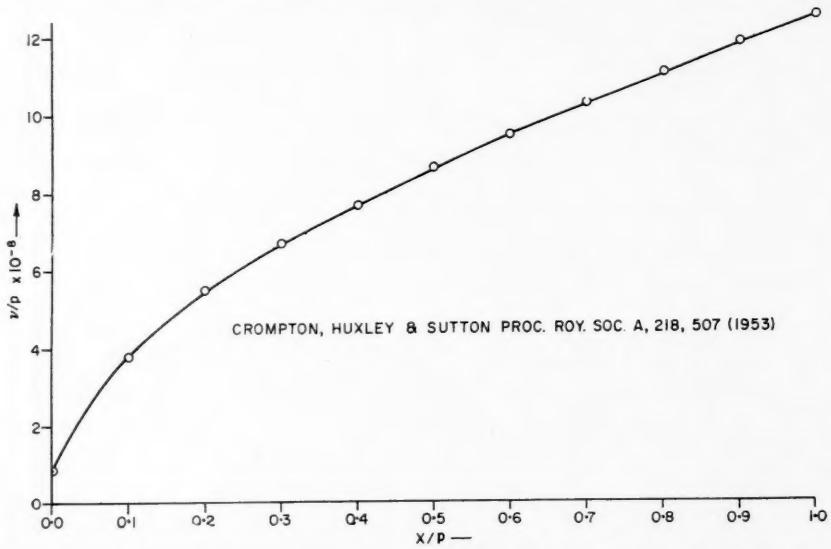


Fig. 7. The variation of electron collisional frequency with applied electric field found by Crompton *et al.* (1953).

Using these results it is possible to calculate the electric field strength necessary to produce absorption of 5 dB at 30 Mc/s, assuming reasonable values of electron density. If the layer responsible for the absorption is again taken as being 10 km in thickness, centred on the 100 km level, and the electron density is taken as 10^4 per cm^3 (a reasonable estimate of the order of magnitude under undisturbed conditions), then the necessary collisional frequency is $\nu = 3{\cdot}8 \times 10^7\ \sec^{-1}$. Taking the Rocket Panel value of pressure at the 100 km level as $p = 4{\cdot}3 \times 10^{-4}$ mm Hg, then $(\nu/p) = 8{\cdot}8 \times 10^{10}$. This figure is well above the experimentally determined range and would require a value of (X/p) much in excess of the maximum value shown in Fig. 7.

If, however, we take the electron density as 5×10^5 per cm^3, the value quoted by Chapman and Little (1957) as the maximum observed by sounding at College during auroral displays, the value of (ν/p) required becomes 18×10^8, giving $(X/p) \simeq 2$, and a value for X of approximately 10^{-3} V/cm.

It is not too difficult to see how field strengths of this order of magnitude can be set up for short periods during particle bombardment of the upper atmosphere.

FAN (1954) has found by laboratory experiments on the excitation of auroral spectra that electrons having energies of a few hundred electron volts are necessary to explain the intensity of the Meinel System of N_2^+ bands as observed in auroral spectra. Since electrons of this order of energy cannot penetrate to the levels of visual aurora, he suggests that these electrons must be accelerated locally, presumably by electric fields set up by the accumulation of incoming charge in localized regions or by the charge separation caused by the fact that incoming electrons and protons will be stopped at different levels in the ionosphere. In fact the field strengths required to explain the N_2^+ band intensities are far in excess of the value of X quoted above. The measurements of CROMPTON *et al.* (1953) lead to a value of about 70 cm for the mean free path of an electron under the ionospheric conditions assumed above. Hence, under a field $X = 10^{-3}$ V/cm an electron would acquire an average energy of 0·06 eV between collisions, whereas FAN's results require energies of several hundred electron volts. However, since the mean free path increases rapidly with height in the ionosphere, it may be that electric fields only an order of magnitude larger than that quoted above will produce the required electron energy in, say, the 120 km region.

Such fields, of course, will be rapidly neutralized by movements of the neighbouring ionization once the primary bombardment ceases or changes its position. The extremely irregular form of the auroral absorption then implies an irregular or patchy pattern of bombardment. The same argument applies if the agency responsible for the absorption is taken as an increase in electron density, the cause of the removal of free electrons being now recombination and attachment, which can also happen with a time constant of a few seconds.

VOL. 14 1959

If the absorption is considered as being due to a combination of increased electron density and increased collisional frequency, the patchiness of bombardment which must be inferred from the irregular nature of the absorption is less than if an increase in N is the only factor. Any change in the bombarding flux will probably change both N and ν in the same direction, leading to a greater change in absorption than if N alone were changed.

To summarize, then, the possibility that local electric fields leading to increases in electron temperature may account for a large part of the observed absorption of type 2 has two main advantages over the earlier theory: (a) the effect can occur at heights of 100 km or even more without invoking enormous electron densities; (b) a more uniform bombardment pattern is permissible in accounting for the irregular appearance of the absorption. This appears to be a desirable feature.

Type 3 (*high-latitude*) *absorption*

An explanation of type 3 absorption must take account of several remarkable features:

(1) The apparent lack of correlation between the occurrence of this absorption and geomagnetic activity;

(2) The fact that the absorption sets in within a few hours of the onset of major solar flares, and has a total duration of 2–3 days;

(3) The large diurnal variation of the intensity of absorption, the ratio between daytime and night-time absorption being of the order of 10 to 1;

(4) The fact that the absorption appears to be more strictly controlled in latitude than the auroral absorption described above.

This last feature forces us to the conclusion that the primary cause of the absorption is corpuscular radiation of some kind, and that it is more strongly influenced by the geomagnetic field than the agency responsible for type 2 absorption. Chapman and Little (1957) and Bailey (1957) have both described absorption phenomena of a very similar type occurring at high latitudes, and have offered explanations. Chapman and Little postulate a continuous bombardment of the upper atmosphere by electrons of energies between 10–100 keV, the diurnal variation of absorption being due to greater attachment of free electrons by night to form negative ions. Considerable evidence to support this view comes from the rocket experiments of Meredith *et al.* (1955), who discovered the existence of soft radiation at heights above 50 km in the auroral zone. This was interpreted as being Bremsstrahlung radiation caused by the impact at higher levels of electrons having energies between 10–100 keV.

Bailey, on the other hand, suggests that the primary cause of the absorption is the arrival, within a few hours of the solar flare, of a large number of solar ions having low electron affinity, which will penetrate to low levels in the ionosphere and will then recombine very slowly, perhaps over a period of days. The large difference between day and night absorption is again taken to be due to negative ion formation at night. This theory has many attractive features, but it is open to certain objections which will not be enumerated here.

There is a certain amount of evidence about the height of the region responsible for this absorption which must also be taken into consideration in any theory. The attenuation of the meteor signals on the VHF forward-scatter circuits suggests that the absorbing region lies below the height at which meteor trails are formed. This height is a function of the geocentric velocity of the meteoric particle which in turn varies with the time of day. It is least in the afternoon when it reaches a minimum value of about 90 km. A second piece of evidence has come from the preliminary results of a rocket firing at Churchill, as reported by Seddon (1957, 1958). This firing took place during daytime on 4 July 1957, while one of the events described above was in progress. The rocket data showed the presence of abnormally high electron densities throughout the lower ionosphere, extending as low as the 55 km level. The conclusion that this type of absorption is caused by increased ionization at very low levels in the ionosphere is further strengthened by the recent report by Landmark (1958) of weak echoes from the 65–70 km height range during daytime polar blackouts in central Norway.

It is thought that the best explanation of this type of absorption will ultimately be found to lie in a mechanism of the Chapman–Little form. The absorption is probably due to the increased ionization produced by solar particle bombardment, and the strong diurnal variation is thought to be caused by night-time negative ion formation. The nature and velocity of the bombarding particles, however, are by no means clear. The suggestion of Chapman and Little that the particles are electrons of energies between 10–100 keV has the considerable merit of being based upon experimental results, but certain objections to it can be raised. The major objection appears to lie in the latitude extent of the absorption events.

The Størmer theory of charged particles moving in the field of a magnetic dipole shows that an electron reaching as far south as Churchill, at geomagnetic latitude 70°, must possess an energy many times greater than the value of 100 keV quoted above. If, on the other hand, the earth's magnetic field is being distorted by, say, a ring-current, so that lower energy particles can reach this latitude, we might expect to see some evidence for the formation of this ring-current in a geomagnetic disturbance on the ground. As pointed out above, no such evidence seems to exist. These facts would appear to indicate that the electrons whose existence was inferred from the rocket experiments are, in fact, secondary in nature. On the other hand it may be worth pointing out that protons having sufficient energy to reach the 55 km level will travel from the sun to the earth in approximately 40 min, and according to the simple Størmer theory will be able to reach geomagnetic latitudes as far south as 70° (see BAILEY, 1957). In order to explain the observations on this basis, we must have a continuous but gradually decaying emission of protons following certain solar flares, the total period of emission lasting 2–3 days. The best explanation of the diurnal effect is probably that of negative ion formation, put forward by CHAPMAN and LITTLE.

In order to explain the lack of geomagnetic activity associated with the incoming protons, the density of the incident proton beam must be low. A rough estimate of this density can be obtained as follows.

VOL. 14 1959

Assuming for simplicity that the incoming protons are sufficiently monoenergetic to be deposited in a layer of 1 km thickness at a height of, say, 65 km, and taking a value of 7 dB as the maximum absorption observed, we have sufficient information to obtain an estimate of the proton density in interplanetary space. Using equation (1) and taking NICOLET's (1953) value for ν at 65 km as $3{\cdot}3 \times 10^7$ sec^{-1}, the required electron density is about $1{\cdot}6 \times 10^5$ per cm^3.

It can easily be shown that if the rate of production of electrons in an ionized region is Q and the recombination coefficient is α, then the electron density when equilibrium is reached will be $(Q/\alpha)^{\frac{1}{2}}$. Hence $Q = \alpha N^2$, and assuming a tentative value of 10^{-8} cm^3 sec^{-1} for α (see CHAPMAN and LITTLE, 1957) the rate of ion production is $Q = 260$ cm^{-3} sec^{-1}; the rate of ion production in the column of 1 km length and 1 cm^2 cross-section is then $2{\cdot}6 \times 10^7$ sec^{-1}.

Protons which penetrate to the 65 km level have an initial energy of about 10^7 eV, and since approximately 30 eV are required for the production of an ion pair each proton will produce approximately 3×10^5 ion pairs. The rate of incidence of protons on the top of the layer is then about 90 cm^{-2} sec^{-1}. Since the velocity of the protons is about $0{\cdot}4 \times 10^{10}$ cm/sec the density of protons in the incident stream is of the order of 2×10^{-8} cm^{-3}. This figure is many orders of magnitude below the densities quoted by CHAPMAN (1950) as giving rise to a moderate magnetic storm, and would explain the lack of any geomagnetic activity associated with the incoming particles leaving the sun for a period of days after a major solar flare.

CONCLUSIONS

A distinction has been drawn between three classes of abnormal VHF absorption phenomena as observed on cosmic noise recordings. The first type, the well-known SID effect, has not been discussed in any detail. The second appears to be closely

associated with the occurrence of visible aurora and geomagnetic disturbance, and an explanation has been offered in terms of an increase in electron temperature at E-region heights caused by local electric fields. The chief advantage in this explanation lies in the fact that it enables the source of the absorption to be located at heights close to those at which visible aurora is most frequently found, without having to invoke enormous electron densities. Evidence in support of this theory can come from rocket measurements of electron density at times when auroral absorption is present. If the measured density is not high enough to explain the measured absorption then some mechanism, such as that outlined above, must be causing an increase in effective collisional frequency.

The third type of absorption phenomenon seems to be strongly correlated with the occurrence of major solar flares, and rocket measurements appear to give definite evidence that the direct cause of the absorption is excess ionization at very low levels in the ionosphere. A possible explanation of this ionization has been suggested in terms of an incoming proton stream, though many factors have been ignored in this suggestion. Not too much can be expected from rocket experiments in the way of identifying the primary particles, since their number is very small in comparison with the number of ions they produce (in the calculation above each proton produces 3×10^5 ion pairs). In this connexion spectroscopic evidence from measurement of the night-sky radiation during periods of strong absorption of this type might prove extremely valuable.

Acknowledgements—The authors wish to acknowledge the invaluable help they have received from discussions with colleagues in the Radio Physics Laboratory, and especially with Dr. P. A. Forsyth and Dr. T. R. Hartz.

The assistance received from Mr. G. B. Craig in the construction and installation of the cosmic noise equipment at both sites is also gratefully acknowledged, as is the work of the DRTE IGY Group in operating the cosmic noise equipment at Churchill.

The authors wish also to acknowledge their indebtedness to the numerous people who at various times and places have contributed to the operation of the VHF scatter circuits.

References

Appleton E. V. and Piggott W. R.	1954	*J. Atmosph. Terr. Phys.* **5,** 141.
Bailey D. K.	1957	*J. Geophys. Res.* **62,** 431.
Chapman S.	1950	*J. Geophys. Res.* **55,** 361.
Chapman S. and Little C. G.	1957	*J. Atmosph. Terr. Phys.* **10,** 20.
Collins C. and Forsyth P. A.	1958	*J. Atmosph. Terr. Phys.* In press.
Crompton R. W., Huxley L. G. H. and Sutton D. J.	1953	*Proc. Roy. Soc.* A **218,** 507.
Currie B. W., Forsyth P. A. and Vawter F. E.	1953	*J. Geophys. Res.* **58,** 179.
Fan C. Y.	1954	*Astrophys. J.* **119,** 294.
Forsyth P. A. and Vogan E. L.	1957	*J. Atmosph. Terr. Phys.* **10,** 215.
Heppner J. P., Byrne E. C. and Belon A. E.	1952	*J. Geophys. Res.* **57,** 121.
Landmark B.	1958	*J. Atmosph. Terr. Phys.* **12,** 79.
Little C. G.	1954	*Proc. Inst. Radio Engrs., N.Y.* **42,** 1700.
Little C. G. and Leinbach H.	1958	*Proc. Inst. Radio Engrs., N.Y.* **46,** 334.

Machin K. E., Ryle M. and Vonberg D. D.	1952	*Proc. Instn. Elect. Engrs.* Pt. III, **99,** 127.
Meredith L. H., Gottlieb M. B. and Van Allen J. A.	1955	*Phys. Rev.* **97,** 201.
Mitra A. P. and Shain C. A.	1953	*J. Atmosph. Terr. Phys.* **4,** 204.
Nicolet M.	1953	*J. Atmosph. Terr. Phys.* **3,** 200.
Nielsen R. A. and Bradbury N. E.	1937	*Phys. Rev.* **51,** 69.
Seddon J. C.	1957	*URSI Twelfth General Assembly,* Boulder, Colorado.
Seddon J. C.	1958	Private communication.
Shain C. A. and Mitra A. P.	1954	*J. Atmosph. Terr. Phys.* **5,** 316.
Vogan E. L. and Campbell L. L.	1957	*Canad. J. Phys.* **35,** 1176.
Wells H. W.	1947	*Terr. Magn. Atmos. Elect.* **52,** 315.

VOL.
14
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 82 to 93. Pergamon Press Ltd. Printed in Northern Ireland

Oblique echoes from over-dense meteor trails*

L. A. MANNING
Radio Propagation Laboratory, Stanford University, Stanford, California

(*Received* 26 *May* 1958)

Abstract—Ray paths are computed for waves refracted by meteor trails having a Gaussian radial distribution of ionization density. From the spreading of initially parallel rays as they pass through the trail, a measure of reflected signal intensity vs. scattering angle is obtained; the results are presented in polar scattering diagrams, valid in the limit of large trail size. Equivalence theorems are derived relating both intensity and scattering angle for rays incident upon the trail at an arbitrary angle to the intensity and scattering angle for rays incident in the plane normal to the trail. Curves are presented showing the dependence of echo duration on forward-scatter-angle with trail orientation as parameter; it is found that the $\sec^2 \phi$ law developed for under-dense trails applies to over-dense trails only if the plane of propagation contains the trail axis. If not, the effective secant exponent may be as small as 0·3. The theory is compared with MCKINLEY and MCNAMARA's duration measurements. It is found that although the general agreement is satisfactory, the details of their experimental results depend on the way that winds change the trail orientation. The ray theory is also compared with KEITEL's wave theory solution. Unfortunately, he could not get wave solutions for dense trails of age greater than 0·357 per cent of the minimum echo duration. Even so, the ray solutions agree with KEITEL's results for scatter angles up to 155°, thus including all angles available to ground-based stations.

1. INTRODUCTION

IN recent years, a great deal of interest has centered upon the use of meteors to propagate sky wave signals over distances of one-thousand or more kilometers. It was shown by ESHLEMAN and MANNING (1954) that for under-dense meteor trails—those having fewer than 10^{14} electrons/m of trail length—echo duration

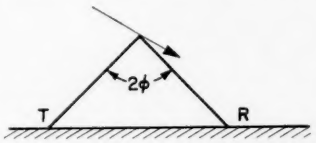


Fig. 1. The geometry of oblique meteor reflection.

should vary with the forward scatter angle ϕ as the secant squared. Fig. 1 illustrates the geometry.

In the over-dense case, however, there is no adequate theory. HINES and FORSYTH have deduced a secant-square duration law for oblique reflection from over-dense trails by neglecting the effect of the ionization beyond the critical density radius (HINES and FORSYTH, 1957). However, in a previous paper (MANNING, 1953) the present author had already shown that the error of this neglect may approach 100 per cent for back-scatter as the echo duration is reached. Since interest in the oblique case is greatest when the duration exceeds that possible for back reflection, severe error may be anticipated from neglecting the ionization beyond the "reflecting radius". Indeed, as time advances, a larger and larger fraction of the ionization will have diffused beyond the reflecting radius, so that

* Jointly supported by the U.S. Army Signal Corps, U.S. Air Force, and the U.S. Navy (Office of Naval Research).

refraction becomes dominant in fixing the scattering pattern. That important deviation from a secant-square law occurs for over-dense trails has been shown experimentally by McKinley and McNamara (1956). In the present paper the properties of such oblique echoes will be computed using the ray-path method. Precise answers can only be expected in the limiting case of large trails, but the approach shares this limitation with all other meteoric echo theories—none exist that can in practice treat more than a limited range of size or density.

2. Ray-paths in Over-dense Meteor Trails

Athough the method of approach to be used is quite general, we shall restrict our discussion to a Gaussian radial distribution of ionization in the meteor trail.

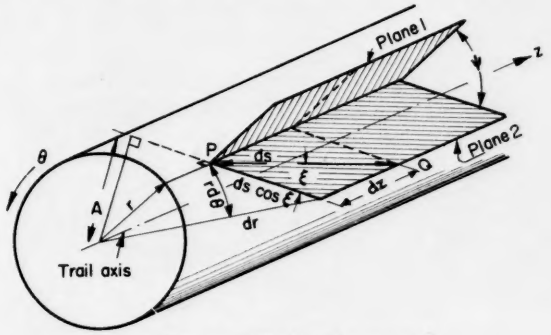


Fig. 2. The co-ordinates of a ray path.

This distribution is perhaps as close to the true distribution as we can now get. So let

$$N = \frac{Q}{4\pi Dt} \exp\,(-r^2/4Dt) \tag{1}$$

where N is the electron volume density, Q the electron line density, D the coefficient of diffusion, r the radial co-ordinate and t is time. We may then express the dielectric constant $\mu^2 = 1 - 81N/f^2$ of trail as

$$\mu^2 = 1 - 1/[\gamma \exp\,(r^2/a\gamma)] \tag{2}$$

where μ is the refractive index, $\gamma = t/t_b$ is the time t divided by the back-scatter duration t_b (at which time $\mu = 0$ at $r = 0$), and $a^2 = 4Dt_b$ is the square of the "trail radius" at the back-scatter duration. The value of t_b is given by

$$t_b = 81Q/4\pi Df^2 \tag{3}$$

where f is the radiofrequency and rationalized M.K.S. units are assumed.

When a ray enters the ionized region about the trail axis, it is bent in two directions. Referring to Fig. 2, let PQ be a differential element of such a ray path. The refractive index at P is determined by the distance r of the ray from the trail axis. The direction of travel of the ray may be described by the two angles, ψ and ξ. The first angle, ψ, is that between the plane containing P and the trail axis (plane 1), and the plane containing PQ and an element parallel to the trail axis (plane 2).

The second angle, ξ, is measured between PQ and a plane perpendicular to the trail axis. If $\psi = 0$, the ray path is in the axial plane, and is governed by Snell's law, $\mu \sin \xi = \text{constant}$. If $\xi = 0$, the ray path is in a plane perpendicular to the trail axis, and is governed by Bouger's rule, $\mu r \sin \psi = \text{constant}$. In general, however, the ray path is non-planar, and the following two relations must be used to find the ray path:

$$\mu \sin \xi = \sin \xi_0 \triangleq S \tag{4}$$

$$\mu r \sin \psi \cos \xi = A \cos \xi_0 \triangleq AC \tag{5}$$

where ξ_0 is the value of ξ outside the ionized region and A is the distance from the trail axis to plane 2.

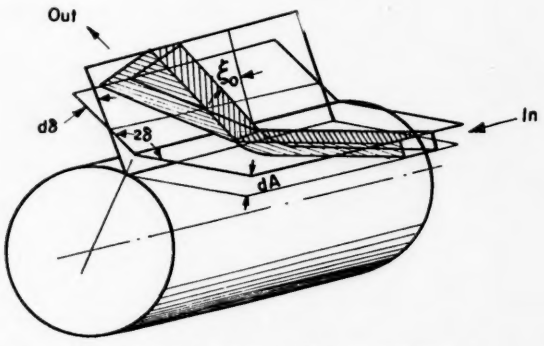


Fig. 3. The refraction and divergence of a narrow beam.

To find the signal intensity refracted from the trail to a distant receiver. consider the spreading imparted to a pencil-like beam of square cross-section. Fig. 3 shows such a beam.

As is evident from equation (4), the beam is not spread in the ξ-direction, but becomes fan-shaped because of spreading in the ψ-direction. At the radius of closest approach the beam is highly pinched and there is a caustic. Since the trail spreads the beam only in one dimension, it will be sufficient if we compute $dA/d2\delta$ as a measure of the scattering properties of the trail. We shall find $d\delta/dA$ by first computing δ, the "half planar-scatter-angle"; this deviation in turn will be found by integrating the change in ψ along the path. Because ξ is of no concern in this integration, we shall eliminate ξ between equations (4) and (5). The result is

$$r\sqrt{(\mu^2 - S^2)} \sin \psi = AC \tag{6}$$

3. Calculation of the Ray Paths

To find the planar-scatter-angle 2δ, we note that

$$\delta = \int_{r_0}^{\infty} \frac{\tan \psi \, dr}{r} \tag{7}$$

where r_0 is the radius of closest approach. Putting equation (6) in equation (7), we have

$$\delta = \int_{r_0}^{\infty} \frac{dr}{r\sqrt{\left(\dfrac{r^2(\mu - S^2)}{A^2C^2} - 1\right)}} \tag{8}$$

Upon substituting μ^2 from equation (2), letting $s = rC$ and $\sigma = \gamma C^2$, we have

$$\delta = \int_{s_0}^{\infty} \frac{ds}{s\sqrt{\left(\frac{\sigma^2\{1 - (1/\sigma)\exp(-s^2/a^2\sigma)\}}{A^2C^2} - 1\right)}} \tag{9}$$

But if $C = 1$, and the entering ray has no axial component, $s = r$, $\sigma = \gamma$, so

$$\delta\Big|_{\xi_0=0} = \int_{r_0}^{\infty} \frac{dr}{r\sqrt{\left[\frac{r^2\{1 - (1/\gamma)\exp(-r^2/a^2\gamma)\}}{A^2} - 1\right]}} \tag{10}$$

an expression of precisely the same form as equation (9). Thus we are led to observe that it is only necessary to compute δ vs. A and γ for $\xi_0 = 0$. For any other value of ξ_0, δ may be found from the following equivalence theorem:

$$\delta(A, \gamma, \xi_0) = \delta(A\cos\xi_0, \gamma\cos^2\xi_0, 0) \tag{11}$$

In the actual calculation, equation (10) is not the most convenient form of the integral for δ, because of the infinite limit. Instead, the variable of integration has been changed to μ, and the relation used is

$$\frac{\pi}{2} - \delta = \int_{\mu_0}^{1} \frac{d\mu}{\mu\sqrt{\left[\frac{\mu^2a^2\gamma\ln\{1/\gamma(1-\mu^2)\}}{A^2} - 1\right]}} \tag{12}$$

Because two parameters must be varied, in addition to the variable of integration, equation (12) is hard to evaluate profusely enough to facilitate numerical differentiation with respect to A. Hence, although critical cases were evaluated from equation (12) using an electronic computer, a simple approximate formula

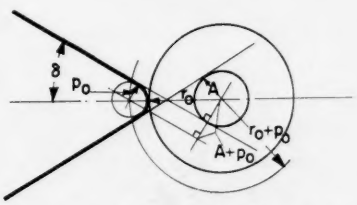


Fig. 4. The radius of curvature approximation for ray deviation.

was used whenever possible. Consider Fig. 4, which applies to propagation in the plane normal to the trail. In this special case from equation (5) we have $\mu(r_0)r_0 = A$, since $\xi = 0$ and $\psi = \pi/2$ at closest approach. Thus for a given γ, or trail age, we may assume r_0 and get A. When a ray is travelling tangentially in a radial distribution of refractive index, the radius of curvature ρ is $\rho = \mu\, dr/d\mu$. Thus at r_0 the ray has a radius of curvature $\rho_0(r_0)$. For radii $r > r_0$, $\rho > \rho_0$. If now we draw circles of radius A and ρ_0 as shown in Fig. 4, the ray path extensions must be tangent to the circle of radius A, and must fall outside the circle of radius ρ_0. Experience shows, however, that for most cases, the deviations are very nearly the same

as if the ray extensions were tangent to the ρ_0 circle. It then follows that, approximately,

$$\sin \delta = \frac{A + \rho_0}{r_0 + \rho_0} \tag{13}$$

where

$$A = r_0 [1 - 1/\{\gamma \exp (r_0^2/a^2\gamma)\}]$$

and

$$\rho_0 = (\gamma/r_0)\{\gamma \exp (r_0^2/a^2\gamma) - 1\}$$

Whenever justified on the basis that equations (12) and (13) give the same results, $d(A/a)/d\delta$ has been computed by evaluating the derivative of equation (13)

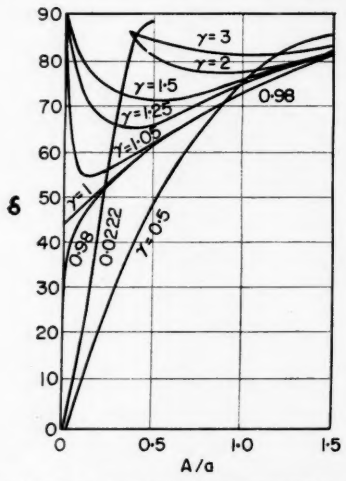


Fig. 5. Dependence of the half planar-scattering-angle δ upon the normalized radius of closest approach A/a.

directly on the computer. However, preliminary slide-rule calculations made with equation (13) were not prohibitive as they were for equation (12).

4. The Scattering Diagrams

In view of the equivalence theorem (11), it is sufficient to compute reradiation patterns for rays lying in a plane perpendicular to the trail axis. Thus we shall interpret this special case before discussing the changes of scale needed to describe incidence at a general angle.

Figure 5 summarizes the behaviour of the half planar-scattering-angle δ as a function of A/a, where a is the radius of the trail at the back-scatter duration, and it is assumed that $\xi = 0$. Recall that A is the radius of closest approach of the undeviated ray path, and γ is t/t_b, time normalized by the back-scatter duration; thus $\gamma > 1$ represents enhanced duration at forward scatter. It will be seen that for $\gamma < 1$, rays are refracted in all directions. However at the back-scatter duration $\gamma = 1$ the minimum planar deviation 2δ is $\pi/2$, and regardless of A none of the rays have a vector component towards the transmitter. When $\gamma > 1$, so that no part of the trail has a negative dielectric constant, the rays directed at the axis

($A = 0$) are neither bent nor reflected; as γ increases the remaining rays experience progressively less maximum deviation.

In Fig. 6 are plotted the normalized polar scattering diagrams for trails of a variety of normalized ages γ, again for the case $\xi = 0$ (propagation in the plane normal to the trail axis.) The radial scale is $d(A/a)/d\delta$ in reciprocal radians. The plots thus represent signal intensity, rather than signal amplitude. Upon

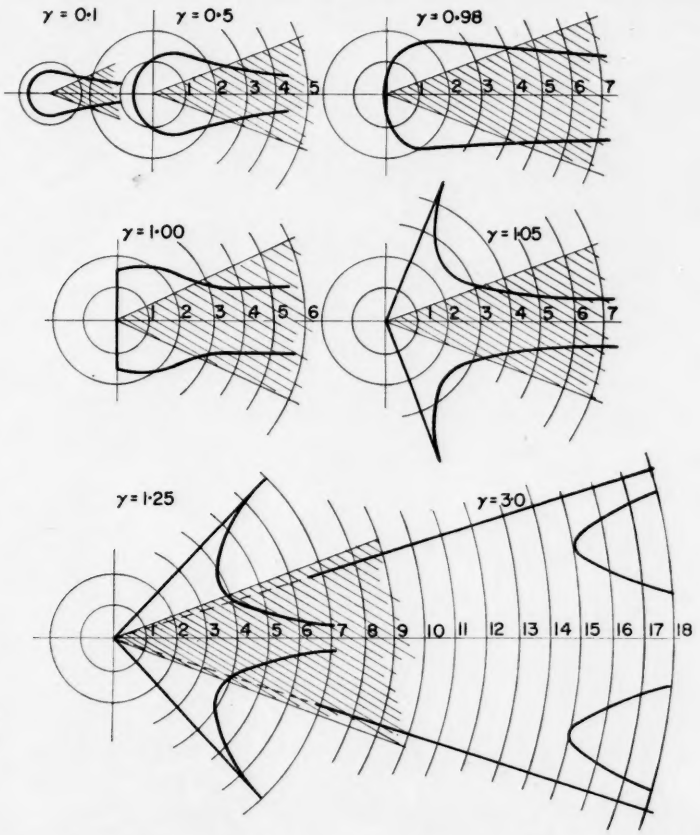


Fig. 6. Normalized polar scattering diagrams for propagation in the plane normal to the trail axis.

inspecting the plots, note first the way in which the back-scatter intensity decreases to zero at $\gamma = 1$. This special case has previously been treated (MANNING, 1953). Next it will be seen that to the extent that ray path calculations are accurate, the duration of the echoes is the same for any transmitter-meteor-receiver angle up to 90°. For larger scattering angles, increased duration is apparent. For $\gamma > 1$, notice that at the moment when the echo drops out the receiver is in a caustic. For such forward-scatter-angles it would be expected that the signal intensity should increase gradually with time to a sharp maximum before dropping out. ALLEN has described oblique path VHF echoes from exceptionally large meteors

VOL. 14 1959

that have these properties (ALLEN 1948). Finally, it may be noted that as γ increases beyond 1, the whole scale of the reradiation pattern becomes larger. The signal intensity then increases roughly as sec δ for deviations less than those corresponding to minimum intensity. Since the plots of Fig. 6 are normalized by a it follows that the magnitude of scattered intensity increases directly with a, and the scattered amplitude as the back-scatter deviation $t_b^{1/4}$.

In general the incident ray will not lie in the plane normal to the trail axis. It is then desirable to determine ξ_0 and δ in terms of more easily measurable quantities. Refer to Fig. 7; it will be seen that ϕ (the half forward-scatter-angle) and β (the

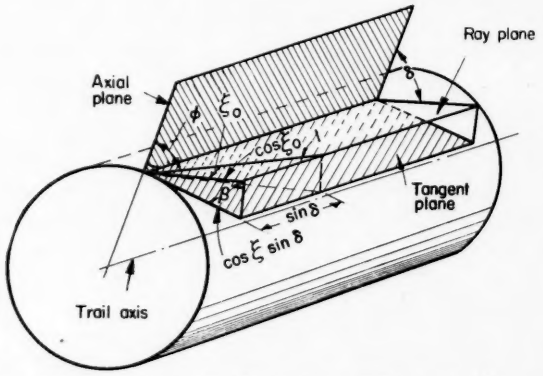


Fig. 7. The co-ordinates of a ray incident upon a trail at an arbitrary angle.

angle between the projection of the original ray direction in the tangent plane, and the intersection of the tangent plane with a plane normal to the trail axis) are more suitable as external measures of the reflection geometry. The meaning of ϕ was already made clear in Fig. 1. The advantage in using β as a co-ordinate is that for random trail orientations, each value of β is equally probable. Note that β as defined here is the complement of that used by ESHLEMAN and MANNING (1954). From Fig. 7 it will now be seen that

$$\tan \xi_0 = \tan \beta \sin \delta \tag{14}$$

and

$$\cos \phi = \cos \xi_0 \cos \delta \tag{15}$$

If β and δ are assumed given, it is a simple matter to compute ξ_0 from equation (14), then ϕ from equation (15). Thus we have a procedure for adapting the normal-plane results of Fig. 5 to arbitrary oblique geometry. Values of cos ξ_0 found using equation (14) are put in the equivalence theorem of equation (11), and the results apply to a forward-scatter-angle 2ϕ found from equation (15).

The polar diagrams of Fig. 6 apply only to the special case of ξ_0 equal to zero. An equivalence theorem exists relating intensities scattered in other directions. Thus from equation (11) it will be seen that

$$\frac{d\delta}{d(A/a)} = \delta' \cos \xi_0$$

if δ' is $d\delta/d(A/a)$ evaluated for $\xi_0 = 0$, but with A and γ replaced by $A \cos \xi_0$ and $\gamma \cos^2 \xi_0$. Hence the intensity I satisfies the equivalence relation:

$$I(\delta, \gamma, \xi_0) = \sec \xi_0 \, I(\delta, \gamma \cos^2 \xi_0, 0) \tag{16}$$

Equations (14) and (15) are required to interpret the intensity patterns in terms of the forward-scatter-angle 2ϕ, and the trail orientation β. In general, the effect of

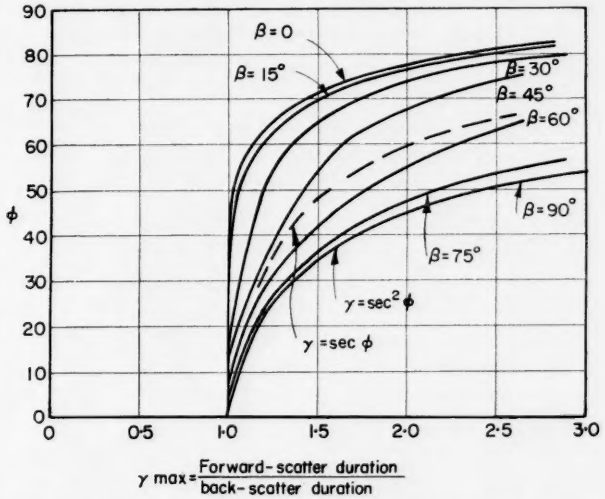


Fig. 8. The relation between half forward-scatter-angle and duration normalized by the back-scatter duration; β expresses the trail orientation, and is zero if the trail is normal to the plane of the transmitter-receiver paths, 90° if the trail is in that plane.

glancing incidence is to increase the duration. For β other than zero, the polar scattering diagrams will not be confined within as narrow a wedge about the forward direction as for incidence normal to the trail axis.

5. Echo Duration and Amplitude

It will be seen immediately from equations (11) or (16) that when ξ_0 is not zero the duration is increased by the factor $\sec^2 \xi_0$, while the effective trail diameter a, and so the intensity, are increased by the factor $\sec \xi_0$. Thus when the trail lies in the plane of propagation the $\sec^2 \phi$ duration increase with obliquity is still to be expected. However, for β not 90°, a much smaller increase results. The actual variation of duration with ϕ and β may be found by noting in Fig. 5 that the echo duration for a particular value of δ, with $\beta = 0$, is determined by the value of γ for that member of the family tangent to the horizontal line δ-equals-a-constant. The duration in the general case is then $\gamma \sec^2 \xi_0$ with ξ_0 found from equation (14) for an assumed β. This duration applies to a half forward-scatter-angle ϕ found from equation (15).

Fig. 8 shows the relation between half forward-scatter-angle and normalized duration computed in the above manner. Note that the curve $\beta = 90°$ is a plot of $\gamma_{max} = \sec^2 \phi$. Thus the duration increase predicted for over-dense trails is less than is expected for all orientations of under-dense trails; there the $\sec^m \phi$ law with

VOL. 14 1959

$m \simeq 2$ derives from the existence of a $t_b \propto \lambda^m$ law of duration versus wavelength regardless of orientation.

Upon inspecting Fig. 8, it is interesting to inquire whether the curves for β not zero may be approximated by laws of the form $\gamma = \sec^m \phi$. It may be concluded that the exponents m must range from 0·3 or 0·4 to represent $\beta = 0$, up to 2·0 for $\beta = 90°$. If the trails are randomly oriented, it may be shown using equation (3) of ESHLEMAN and MANNING (1954), that the median value of β is **45°**. For this β a secant law with exponent about 0·85 is a representative fit.

6. COMPARISON WITH EXPERIMENT

The only controlled experimental test of the dependence of duration on obliquity available to the writer is due to MCKINLEY and MCNAMARA (1956). They operated a 33 Mc/s pulse transmitter at Ottawa and received simultaneously at Ottawa and Scarboro, 337·8 km distant. The echoes studied ranged in duration from roughly 1 to 100 sec. From range data the forward-scatter-angle 2ϕ was found and the value of the exponent m best fitting the relation $\gamma_m = \sec^m \phi$ was presented. Separate determination of m for t_b less than and greater than 3·5 sec yielded $m = 1{\cdot}73$ and $m = 1{\cdot}13$ with probable errors in the means of 0·11 and 0·05. Some tendency for m to decrease with increase in the value of sec ϕ was also noted.

Because of the geometry of the paths in MCKINLEY and MCNAMARA's experiment, specular meteor reflections from linear trails not distorted by atmospheric winds should only have been obtained if β were zero. Thus at first glance it might appear that a theoretical m of 0·3 or 0·4 should be used in the comparison. However a number of difficulties arise that we shall now discuss.

First note that MCKINLEY and MCNAMARA break the distribution in two at $t_B = 3{\cdot}5$ sec. They call the two groups "under-dense" and "over-dense". However the theoretical division between under-dense and over-dense trails occurs in the bridging region where the under-dense and over-dense duration formulae merge. For back-scatter these equations are

$$t_{ud} = \lambda^2/16\pi^2 D \tag{17}$$

and

$$t_{od} = 0{\cdot}885 \times 10^{-14}\, Q\lambda^2/4\pi^3 D \tag{18}$$

where D is the coefficient of diffusion. The bridging region occurs approximately where $0{\cdot}885 \times 10^{-14}\, Q = 1$, so we see the critical duration is about

$$t_{crit} = \lambda^2/140 D$$

Taking $\gamma = 9{\cdot}1$ m and $D = 3$ m^2/sec, we find t_{crit} is about 0·2 sec. Hence we must assume that most of MCKINLEY and MCNAMARA's meteors are over-dense. However, we expect a gradual transition from the $\sec^2 \phi$ behaviour of under-dense trails to the behaviour predicted on the basis of ray optics. A density well above the critical value, and in the region of MCKINLEY and NCNAMARA's so-called "over-dense" sample, seems appropriate for the comparison with theory.

In MCKINLEY and MCNAMARA's "over-dense" sample, however, m was found to be 1·13. This is much greater than the theoretical value for $\beta = 0$. Now another influence on the trail must be considered. It was shown some time ago by MCKINLEY

VOL
14
195

and MILLMAN (1947) that in the case of those meteoric echoes enduring over 5–10 sec the aspect sensitivity of detection is lost. The author has developed the detailed theory of the loss of aspect sensitivity (MANNING, 1957) and finds it is a direct result of distortion of the originally-straight meteor trails by winds. We see therefore, that we must no longer assume β to be zero, since after 5–10 sec some part of the trail will be turned to present a value of β as large as 90°. Since the echo duration will be fixed by that part of the trail turned through the most favourable angle, we see that the mean value of m should lie somewhere between 0·3 and 2·0. Individual observations greater than $m = 2$ can be expected to result when turning of the trail occasionally makes back-scatter detection impossible. Clearly an accurate check of the present theory is not possible based on MCKINLEY and MCNAMARA'S data. However the fact that their mean m of 1·13 is very nearly the average of 0·3 and 2·0 is encouraging. And, as they point out, in view of the large percentage of over-dense trails in their "under-dense" sample, their mean m of 1·73 does not mean a failure of the secant squared law for under-dense trails.

7. COMPARISON WITH WAVE THEORY SOLUTIONS

Attempts to determine the applicability of the present ray theory solutions by comparison with exact wave solutions are hampered by the extreme difficulty of obtaining wave solutions for dense trails of appreciable age. KEITEL has carried the wave approach about to the limit (KEITEL, 1955) but his most ambitious example, calculated on a large electronic computer, carries the case with line density $Q = 10^{17}$ electrons/m only to a time equal to 0·357 per cent of the back-scatter duration. At this time the trail radius is only $1/\pi$ wavelengths. Such durations allow much less time for expansion than do the values of normalized time γ of one or more for which the ray theory has been developed. Thus no valid comparison of over-dense echo behaviour with exact calculations is possible in a region where both may be expected to apply. It is none-the-less interesting to note how good the ray results are even for KEITEL'S very small trails.

KEITEL expresses his results in terms of a reflection coefficient defined by

$$\rho = \sqrt{\left(\frac{\pi k R_0}{2}\right)}\left(\frac{\text{scattered field at } R_0}{\text{incident field}}\right) \tag{19}$$

where $k = 2\pi/\lambda$ and R_0 is the distance from the meteor to the receiver. Our results are presented in terms of the angle $d\delta$ through which energy confined between parallel rays $d(A/a)$ apart is spread. In terms of $dA/d\delta$ the second bracketed quantity in equation (19) becomes

$$\frac{\text{scattered field}}{\text{incident field}} = \sqrt{\{(1/2R_0)dA/d\delta\}} \tag{20}$$

Now $a^2 = 4Dt_b$, so $a = \lambda[0{\cdot}885 \times 10^{-14}\, Q/\pi^3]^{1/2}$
and

$$\rho = 5{\cdot}13\sqrt{\{d(A/a)/d\delta\}} \tag{21}$$

In Fig. 9, KEITEL'S wave solution for $k(a\gamma^{1/2}) = 2$, $Q = 10^{17}$ and $\beta = 0$ has been compared with the ray-path calculation for $\gamma = 0{\cdot}00357$ to which it corresponds. It will be noted that the agreement is excellent for forward-scatter-angles

2δ up to about 155°. In the cross-hatched region, however, the exact polar scattering diagram has a broad peak of finite amplitude; the ray solution has a narrow infinite peak. The discrepancy in this region must be expected in view of the very small γ that KEITEL used. For such a small trail, the exact ionization distribution is not of great importance, and as KEITEL shows, the shape of the true overall scattering diagram is similar to the wave solution for a metal cylinder. However,

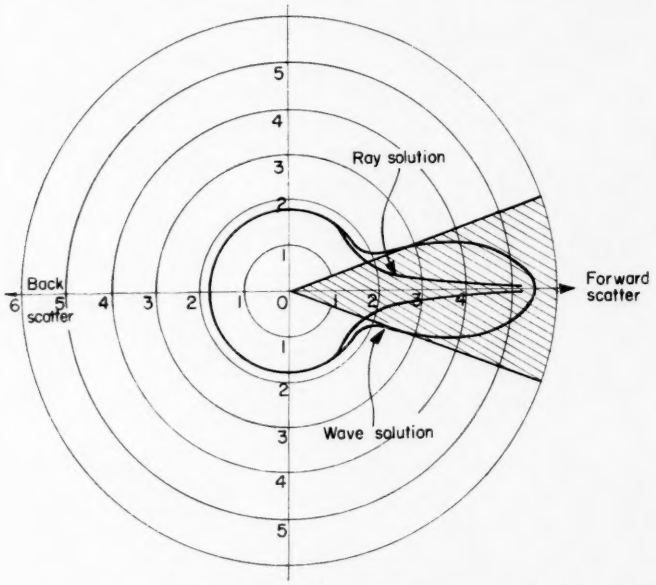


Fig. 9. Comparison of ray solution with KEITEL's wave solution for a line density of 10^{17} electrons/m, and $\gamma = 0{\cdot}00357$.

for $\delta < 80°$, the ray theory gives a better fit than the metal cylinder even with γ only 0·00357. It is expected that for γ's in the pertinent range near unity, the metal cylinder approximation will no longer be valid, and the ray treatment will be useful over an extended range of forward-scatter angles. It may be further noted that values of δ greater than 80° are not possible for transmission between ground-based equipments. Thus the cross-hatched area of Fig. 9 is excluded on practical grounds.

Although ray theory makes no distinction between incident waves with electric polarization along or across the trail, wave theory shows cross-polarization to be more complex. It is therefore expected that for parallel-polarization the ray theory may be applied to smaller trails than for cross-polarization.

8. CONCLUSIONS

Experimental studies of the duration of forward-scattered meteor echoes from over-dense trails have shown that the $\sec^2 \phi$ duration increase with obliquity predicted for under-dense trails does not hold. No theory of long-duration echoes has existed capable of discussing this question. The present theory, applicable in the limiting case of large Gaussian ionization trails, demonstrates that the effective

exponent of sec ϕ for forward-scatter may vary from about 0·3 to 2·0 depending upon the orientation of the trail. In estimating practical duration increases, it is necessary to take into account in addition the distortion and change of orientation of the trail due to winds. Thus the mean practical exponent may be somewhat above the median expected for random trail orientation; the most favourably oriented parts of the trail will determine the observed duration.

Comparison of the theoretical polar scattering diagrams with wave theory calculations is unsatisfactory because no existing wave solutions do more than dent the over-dense trail problem. However, the ray solutions compare closely with the most ambitious wave solutions over the experimentally observable range of angles despite the small trail radius of the examples.

An equivalence theorem has been found relating the ray solutions for scattering at an arbitrary angle with the trail axis to the scattering in the plane normal to the axis. If the plane of propagation contains the trail axis, the $\sec^2 \phi$ duration law holds even for over-dense trails. As the plane of propagation deviates farther and farther from the axis the predicted duration increase becomes smaller and smaller, but never vanishes entirely.

References

Allen E. W., Jr.	1948	*Proc. Inst. Radio Engrs., N.Y.* **36,** 346.
Eshleman V. R. and Manning L. A.	1954	*Proc. Inst. Radio Engrs., N.Y.* **42,** 530.
Hines C. O. and Forsyth P. A.	1957	*Canad. J. Phys.* **35,** 1033.
Keitel G. H.	1955	*Proc. Inst. Radio Engrs., N.Y.* **43,** 1481.
Manning L. A.	1953	*J. Atmosph. Terr. Phys.* **4,** 219.
Manning L. A.	1959	Proceedings of Fifth Meeting, I.C.S.U. Mixed Commission on the Ionosphere. *J. Atmosph. Terr. Phys.* In press.
McKinley D. W. R. and Millman P. M.	1949	*Proc. Inst. Radio Engrs., N.Y.* **37,** 364.
McKinley D. W. R. and McNamara A. G.	1956	*Canad. J. Phys.* **34,** 625.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 94 to 106. Pergamon Press Ltd. Printed in Northern Ireland

Study of horizontal drifts in the *F1*- and *F2*-regions of the ionosphere at Waltair (17°43′N, 83°18′E, mag. lat. 9°30′N)

B. Ramachandra Rao and E. Bhagiratha Rao
Ionospheric Laboratory, Andhra University, Waltair

(*Received* 17 *April* 1958)

Abstract—Results of the analysis of drift measurements made at Waltair for the *F1*- and *F2*-regions over a period of 2 years are presented in this paper. The main features of the diurnal and seasonal variation of drifts in *F2*-region are given in detail. The general pattern of drift movements for this low latitude station is found to be in phase opposition to those observed at high latitudes, thus confirming *F2*-region drift theory given by Martyn (1955).

1. Introduction

A number of radio methods are now available for the study of horizontal movements in the ionosphere. In a recent paper Briggs and Spencer (1954) have given a brief review of these various methods and the results obtained by these different methods have been discussed. Among the various radio methods available, the spaced-receiver method due to Mitra (1949) is widely used by several investigators. From a fairly exhaustive review referred to earlier, it appears that a systematic study of drifts in the *E*- and *F2*-regions of the ionosphere by the spaced receiver method is in progress at Cambridge, England (52°13′N, 00°6′E and mag. lat. 54°44′*N*) over a period of years. Though ionospheric drifts have been studied in detail at some higher latitude stations it is surprising that very little drift data is available for low latitude stations. Further, *F1*-region drifts have not been studied systematically at any place. It is for these reasons we have taken up the systematic study of drifts in the *F1*- and *F2*-regions of the ionosphere at Waltair (17°43′N, 83°18′E and mag. lat. 9°30′N), which is a fairly low latitude station, during the last 4 years. The spaced-receiver method due to Mitra is employed in this investigation after slightly modifying the method of recording. The complete details about the experimental set-up and the preliminary results of drift measurements in *E*- and *F2*-regions have been reported by Rao, B. R. *et al.* (1956). Systematic measurements of *F1*- and *F2*-region drifts with particular reference to the diurnal behaviour of *F2* region drifts have been under investigation since January 1956 and the results of two years of extensive study are presented in this communication.

2. Results and Discussion

For the purpose of investigating the seasonal variation, *F2*-region drift measurements were made on a frequency of 6·0 Mc/sec at 0900, 1200 and 1500 hours IST for 3–4 days each week. Diurnal study of *F2*-region drifts has been made by taking observations at hourly and sometimes half-hourly intervals for 2–3 "quiet days" each month for the period February 1956 to May 1957. During the remaining period covering the IGY, observations were taken for at least 3 consecutive days

including some World Days. The total number of *F2*-drift records taking during the last 2 years is about 2700 out of which 1500 are diurnal records and the rest were taken at 0900, 1200 and 1500 hours. Of these about 1000 diurnal records and 800 records taken at the three specified hours showed fading patterns suitable for drift measurement. For the diurnal study of drifts in the *F2*-region, the operating frequency for most of the daytime is nearly 6·0 Mc/sec. The equivalent height of reflection for this frequency during daytime is found to be generally between 280 km and 320 km. In order to determine the drifts at about the same

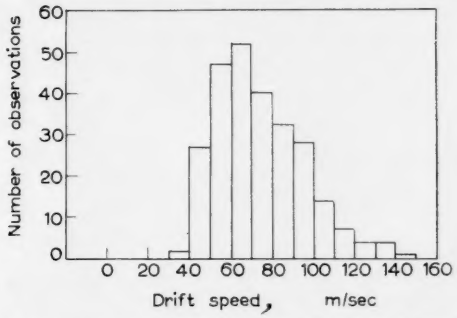


Fig. 1. Histogram showing the variation of drift speeds in the *F1*-region during 1956 and 1957.

VOL. 14 1959

level the night-time operating frequency was suitably decreased in order to maintain the equivalent height of reflection at night-time to be the same as the average equivalent height of reflection during daytime. But care was taken to see that the *F2*-region operating frequency is always sufficiently higher than the critical frequency of the lower *F1*- and E_s-regions during daytime and E_s-region alone at night. To satisfy these conditions the operating frequency during nights was suitably adjusted and was found to be between 3·0 Mc/s and 5·0 Mc/s.

The *F1*-region drift measurements for seasonal study were made at 0900, 1200 and 1500 hours IST for 3 days in a week during the years 1956 and 1957. Out of about 600 *F1*-drift records taken during the last 2 years at the specified hours, only half of them were found to be suitable for drift measurements. It is found to be necessary to vary the *F1* operating frequency occasionally in order to avoid partial reflections from the lower *E*- and E_s-regions. Observations on *F1*-region were never taken when partial reflections from *E*- and E_s-region were observed. The frequency employed for *F1*-drift study usually lies in the range of 4·2 Mc/s to 4·8 Mc/s and the virtual heights vary from about 200 km to 240 km.

2.1. *Magnitude of drifts in the F1-region*

A histogram of *F1*-drift speeds is shown in Fig. 1. This histogram was prepared by taking an equal number of observations for the three times 0900, 1200 and 1500 hours. From the histogram it can be seen that the speed of *F1*-drifts generally lies between 30 m/sec and 150 m/sec. It is interesting to note that the most probable range of drift speed for *F1*-region is 40 m/sec to 100 m/sec as nearly seven-eighths of the observations lie in this range. It appears from the histogram that the most frequent and the average drift speed lies between 60 and 70 m/sec. An interesting feature of the histogram is that it shows a steep rise in the frequency of occurrence

of the drift speed in the range 40 m/sec to 70 m/sec and a gradual decrease in the frequency of occurrence of the speeds above 80 m/sec. A study of drift speeds taken at the three times 0900, 1200 and 1500 hours IST has shown that there is no systematic diurnal variation of the drift speed except that the noon time values are higher by about 10 per cent. Comparison of drift speeds obtained in each season separately during 1956 and 1957 has shown that there are no marked variations from season to season.

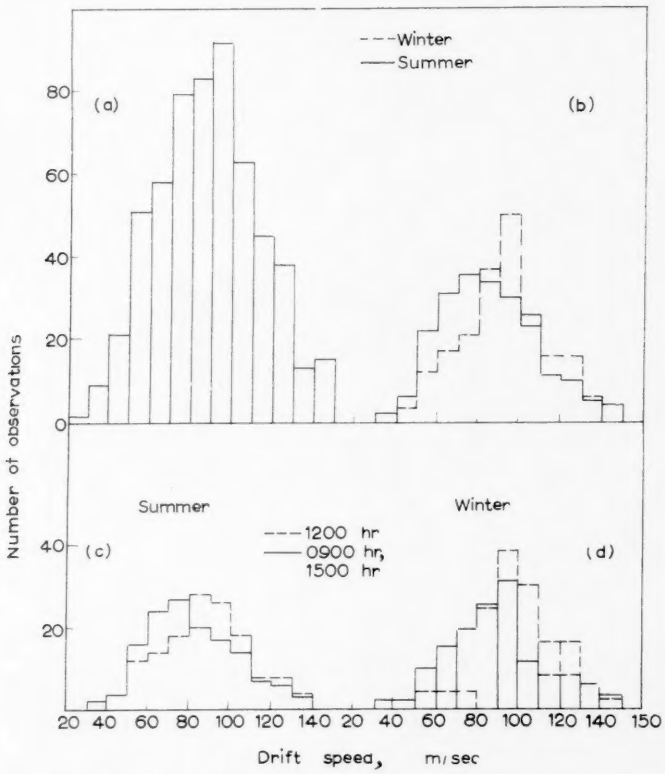


Fig. 2. Histogram showing the variation of drift speeds in the *F2*-region during 1956 and 1957.

The only other observations on *F1*-drifts available for comparison are those of Yerg (1955) on 4·57 Mc/s by spaced-receiver method and Toman (1955) on 3·5 Mc/s by a somewhat different method. Yerg has reported a velocity of the order of 50 m/sec and has also concluded from his studies that there is no seasonal variation in *F1*-drift speeds. Toman's observations at Cambridge (Mass.) refer to the *F1*-region only in summer and drift speeds of the order of 70 m/sec were obtained. The drift speeds reported by these authors are in fairly good agreement with the results of our investigations in the *F1*-region.

2.2. *Magnitude of drifts in the F2-region*

Histograms showing the diurnal and seasonal effects on the drift speeds of the *F2*-region are presented in Fig. 2. Fig. 2(a) is the histogram of the number of

observations vs. the drift speed obtained by collecting and plotting equal number of observations taken at all the three times 0900, 1200 and 1500 hours IST and during all seasons. It can be seen from the histogram that the *F2*-drifts are confined to a range of 30 to 150 m/sec and exhibit a maximum frequency of occurrence in the range of 90 to 100 m/sec. Hence the most frequent drift speed may be taken to be about 95 m/sec. It can also be seen that over half the number of observations are confined to a range of 70 to 110 m/sec, while more than three-quarters are in the range of 60 to 120 m/sec. The histogram shows a gradual increase and a rapid fall before and after the maximum value of the frequency of occurrence. It may be noted that the feature is quite the reverse of the behaviour observed for the histogram of *F1*-drifts.

With a view to bringing out vividly the difference in the distribution of *F2*-drifts between summer and winter seasons, histograms of the observations taken during the two seasons are plotted separately and are shown superimposed on each other in Fig. 2(b). In plotting these histograms equal number of observations are taken for each hour and season. Comparison of the histograms in this figure shows that the most frequent drift speed has increased from 75 m/sec in summer to 95 m/sec in winter. It is interesting to note that there are more observations having higher speeds in winter than in summer. It may therefore be concluded that the winter drift speeds are in general higher than the summer values.

VOL. 14 1959

A study of the average drift speeds at each of three hours of observation for each season revealed the interesting fact that the noontime values are, in general, higher than the values during the forenoon and afternoon hours which are very nearly of the same order. This difference is clearly shown in Fig. 2(c) and 2(d) in which the histograms of noon observations and observations for other hours are shown separately for the summer and winter seasons. In drawing these histograms, equal numbers of observations are taken for the noontime and other hours. It will be clearly seen from these histograms that there are more observations showing higher speeds during noontime as compared to the other two hours. Another interesting feature which is noticed only in the summer observations is the fact that the most probable velocity for noontime observations is higher than the value for the forenoon and afternoon observations.

In order to study the diurnal variation of drift speeds in greater detail hourly or half hourly observations of drift speeds obtained during some days in each season are grouped together and the median values for each period of 1 or $\frac{1}{2}$ hr are arrived at. Although a plot of the drift speeds thus obtained against the time of the day shows clearly the general trend of diurnal variation in the drift speeds, there are some irregular short time fluctuations noticed in these curves. In order to eliminate these irregularities and to obtain a smooth curve representing very nearly the true diurnal variation, the following procedure of plotting the same observations is adopted. For this purpose average of equal numbers of observations taken within 1 hr on either side of a particular hour is taken as representative of the drift speed at that hour. Curves thus obtained for the variation of drift speed with the time of the day are shown in Fig. 3. From a careful study of these curves we arrive at the following conclusions:

(1) In the curves for all the seasons except winter the drift speed is found to

reach maximum values and minimum values twice a day in general. The winter drift speeds do not show any marked variation. The maxima and minima are quite prominent, particularly for the spring season. Unlike the other seasons it is interesting to note that winter readings do not show any prominent fluctuations particularly for the winter 1957–1958.

(2) It is found that the maximum values of drift speed are attained at about noon and midnight, while the minimum values of the drift speed are attained near

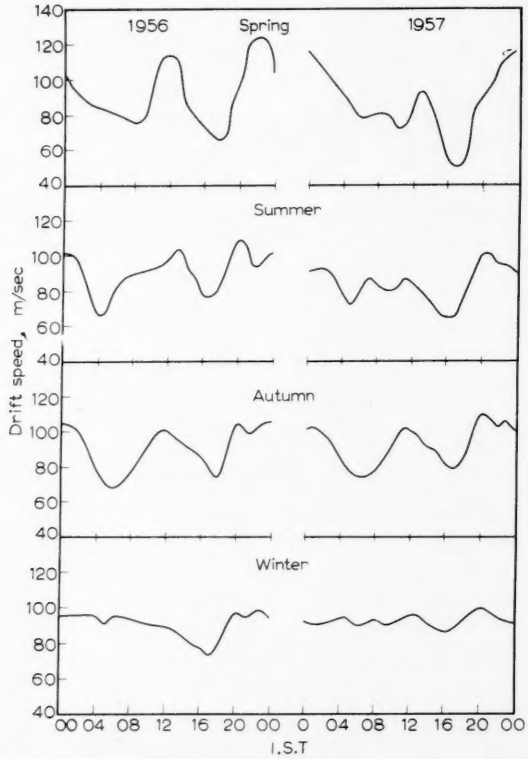


Fig. 3. Diurnal variation of the magnitude of *F2*-drifts in all the seasons for the years 1956 and 1957.

sunrise and sunset hours. There are seasonal variations in the times of occurrence of these maxima and minima. The time interval between the noon maximum and evening minimum appears to be lesser in summer season and more in spring and autumn seasons. The forenoon minimum occurring in the spring season is not as prominent as that in the summer and autumn seasons, when it occurs as a sharp dip in the curves.

(3) In all seasons except in the spring there occurs a subsidiary maxima at about 2000 hours. In spring season the subsidiary maxima appear to have been obscured by the rapid rise of drift speed from the evening minimum to midnight maximum in the spring season. An explanation for the observation of this subsidiary maximum is found in the rapid rise of virtual height of reflection observed

after sunset. It has been observed more or less commonly at this latitude that there is a rapid increase of virtual height of reflection by about 100 km around this time, even though the operating frequency is reduced by some 3·0 Mc/sec. It may be noted that this phenomenon is a common feature observed at several equatorial latitudes and has been reported by WELLS and STANTON (1938) at Huyancayo and OSBORNE (1951) at Singapore. On account of the actual increase of the virtual height of *F2*-region around this time the drift speed measured during

Table 1. *F1*-Drift directions for 1956 and 1957

Season	0900	1200	1500
Spring	SW	SW	NE
Summer	NE	SW	NE, NW*
Autumn	NE, NW*	SW	SW, NE*
Winter	SW	SW	SW

this period actually corresponds to a higher level having higher drift speeds, and hence the subsidiary maxima is observed. Another subsidiary maximum is observed in summer records before noon but no plausible explanation could be given for this observation.

(4) Results obtained in the spring season show that in general the night-time drift speeds are higher than the daytime values. For the 2 years the midnight maxima in this season are found to be higher than the noon maxima. The lowest drift speed also occurs for this season in the evening hours. Thus we find that there is a rapid increase of drift speed from evening to midnight only in the spring season. This change in the drift speed has a magnitude of the order of 60 m/sec while in the other seasons this variation is not more than 30 m/sec.

(5) Though the general trend of speed variation remains unchanged from 1956 to 1957, significant differences are also apparent. The differences between the midnight and noon maxima are more pronounced for the spring and summer seasons during 1958 as compared to the previous year.

2.3. *Seasonal changes in F1 drift directions*

Considerable data is available for drift measurements of *F1*-region during daylight hours of 0900, 1200 and 1500 hours. Polar diagrams were drawn for all the observation of each season for each hour and the main drift directions as obtained from the polar diagrams are shown in Table 1. Sometimes a second prominent direction also is indicated and these are also shown in the table with an asterisk. It is interesting to note that there is a systematic trend in the drift directions. The drift direction is SW for all seasons at noontime. The drift direction at 0900 hours is towards SW during spring and winter, and NW during summer and autumn. At 1500 hours the drift is towards SW during autumn and winter while it is towards NE during spring and summer. Thus the duration for which the drift direction is towards SW is larger for winter months and is smaller

Table 2. Results of harmonic analysis of *F2*-region drifts for every month (t = local solar angle)

Month	N–S component	E–W component	Phase difference
Feb. 1956	$-18{\cdot}8 + 87{\cdot}5 \sin (t + 132°)$	$+6{\cdot}7 + 90{\cdot}5 \sin (t + 60°)$	72°
Feb. 1957	$+6.5 + 75{\cdot}5 \sin (t + 138°)$	$+18{\cdot}8 + 87{\cdot}9 \sin (t + 59°)$	79°
March 1956	$-4{\cdot}4 + 67{\cdot}1 \sin (t + 151°)$	$+8{\cdot}6 + 97{\cdot}1 \sin (t + 87°)$	64°
March 1957	$-7{\cdot}3 + 56{\cdot}0 \sin (t + 143°)$	$+14{\cdot}1 + 87{\cdot}0 \sin (t + 52°)$	91°
April 1956	$+1{\cdot}9 + 42{\cdot}4 \sin (t + 130°)$	$+8{\cdot}2 + 74{\cdot}9 \sin (t + 51°)$	79°
April 1957	$+1{\cdot}3 + 66{\cdot}5 \sin (t + 126°)$	$+2{\cdot}2 + 74{\cdot}0 \sin (t + 47°)$	79°
May 1956	$+16{\cdot}6 + 30{\cdot}3 \sin (t + 226°)$	$+38{\cdot}2 + 68{\cdot}6 \sin (t + 66°)$	160°
May 1957	$+1{\cdot}3 + 74{\cdot}3 \sin (t + 138°)$	$-5{\cdot}1 + 91{\cdot}8 \sin (t + 44°)$	94°
June 1956	$-10{\cdot}1 + 56{\cdot}8 \sin (t + 106°)$	$-0{\cdot}9 + 76{\cdot}0 \sin (t + 47°)$	59°
June 1957	$-6{\cdot}6 + 78{\cdot}7 \sin (t + 127°)$	$+1{\cdot}3 + 79{\cdot}6 \sin (t + 39°)$	88°
July 1956	$+3{\cdot}6 + 88{\cdot}2 \sin (t + 149°)$	$+24{\cdot}1 + 82{\cdot}8 \sin (t + 71°)$	78°
July 1957	$+6{\cdot}4 + 77{\cdot}6 \sin (t + 136°)$	$+1{\cdot}4 + 83{\cdot}6 \sin (t + 57°)$	79°
Aug. 1956	$+1{\cdot}2 + 73{\cdot}5 \sin (t + 158°)$	$+4{\cdot}2 + 98{\cdot}2 \sin (t + 72°)$	86°
Aug. 1957	$-9{\cdot}3 + 79{\cdot}5 \sin (t + 140°)$	$+5{\cdot}0 + 89{\cdot}1 \sin (t + 53°)$	87°
Sept. 1956	$+2{\cdot}4 + 60{\cdot}8 \sin (t + 138°)$	$-3{\cdot}2 + 94{\cdot}6 \sin (t + 56°)$	82°
Sept. 1957	$-9{\cdot}8 + 69{\cdot}0 \sin (t + 142°)$	$-8{\cdot}2 + 78{\cdot}5 \sin (t + 45°)$	97°
Oct. 1956	$-0{\cdot}9 + 32{\cdot}5 \sin (t + 164°)$	$-13{\cdot}9 + 69{\cdot}2 \sin (t + 75°)$	89°
Oct. 1957	$+3{\cdot}2 + 75{\cdot}9 \sin (t + 125°)$	$+13{\cdot}2 + 89{\cdot}2 \sin (t + 88°)$	37°
Nov. 1956	$-0{\cdot}3 + 62{\cdot}4 \sin (t + 182°)$	$-1{\cdot}3 + 78{\cdot}6 \sin (t + 96°)$	88°
Nov. 1957	$-11{\cdot}8 + 66{\cdot}9 \sin (t + 165°)$	$+8{\cdot}0 + 91{\cdot}0 \sin (t + 90°)$	75°
Dec. 1956	$-25{\cdot}3 + 55{\cdot}0 \sin (t + 125°)$	$+11{\cdot}2 + 77{\cdot}0 \sin (t + 75°)$	50°
Dec. 1957	$-17{\cdot}8 + 71{\cdot}0 \sin (t + 151°)$	$-6.8 + 93{\cdot}5 \sin (t + 55°)$	96°
Jan. 1957	$-4{\cdot}6 + 76{\cdot}4 \sin (t + 133°)$	$+16{\cdot}8 + 90{\cdot}3 \sin (t + 58°)$	75°
Jan. 1958	$+4{\cdot}7 + 71{\cdot}2 \sin (t + 143°)$	$+3{\cdot}7 + 88{\cdot}9 \sin (t + 82°)$	61°

Table 3. Results of harmonic analysis of *F2*-region drifts for every season (t = local solar hour angle)

Spring 1956 v(N–S)	$-1{\cdot}6 + 77{\cdot}3 \sin (t + 130°) + 7{\cdot}7 \sin (2t + 270°)$
Spring 1956 v(E–W)	$+16{\cdot}6 + 80{\cdot}0 \sin (t + 73°) + 12{\cdot}8 \sin (2t + 221°)$
Summer 1956 v(N–S)	$-5{\cdot}9 + 74{\cdot}6 \sin (t + 121°) + 6{\cdot}8 \sin (2t + 270°)$
Summer 1956 v(E–W)	$+12{\cdot}2 + 77{\cdot}5 \sin (t + 54°) + 10{\cdot}0 \sin (2t + 190°)$
Autumn 1956 v(N–S)	$+0{\cdot}3 + 65{\cdot}8 \sin (t + 139°) + 13{\cdot}0 \sin (2t + 278°)$
Autumn 1956 v(E–W)	$-4{\cdot}5 + 88{\cdot}6 \sin (t + 61°) + 6{\cdot}6 \sin (2t + 121°)$
Winter 1956–1957 v(N–S)	$-8{\cdot}8 + 71{\cdot}2 \sin (t + 142°) + 17{\cdot}8 \sin (2t + 194°)$
Winter 1956–1957 v(E–W)	$+18{\cdot}3 + 97{\cdot}0 \sin (t + 76°) + 9{\cdot}1 \sin (2t + 241°)$
Spring 1957 v(N–S)	$-4{\cdot}4 + 68{\cdot}2 \sin (t + 129°) + 5{\cdot}3 \sin (2t + 218°)$
Spring 1957 v(E–W)	$+12{\cdot}1 + 85{\cdot}9 \sin (t + 58°) + 16{\cdot}0 \sin (2t + 18°)$
Summer 1957 v(N–S)	$-0{\cdot}1 + 76{\cdot}0 \sin (t + 134°) + 14{\cdot}4 \sin (2t + 187°)$
Summer 1957 v(E–W)	$+3{\cdot}0 + 86{\cdot}5 \sin (t + 59°) + 4{\cdot}8 \sin (2t + 138°)$
Autumn 1957 v(N–S)	$-3{\cdot}5 + 69{\cdot}2 \sin (t + 132°) + 12{\cdot}8 \sin (2t + 187°)$
Autumn 1957 v(E–W)	$+1{\cdot}7 + 88{\cdot}2 \sin (t + 59°) + 3{\cdot}0 \sin (2t + 210°)$
Winter 1957–1958 v(N–S)	$-13{\cdot}3 + 68{\cdot}0 \sin (t + 148°) + 18{\cdot}9 \sin (2t + 158°)$
Winter 1957–1958 v(E–W)	$+17{\cdot}4 + 93{\cdot}8 \sin (t + 76°) + 35{\cdot}7 \sin (2t + 220°)$

VOL
14
195

for summer months. The duration of SW drift direction in spring and autumn appears to be the same with the difference that the onset of SW direction occurs earlier in spring and later in autumn.

YERG (1955) has observed *F1*-drift directions towards WNW and NNE for summer and towards SW with some scatter in directions during winter. TOMAN (1955) has reported drift direction towards NE for summer months around noontime. It will be however evident that the drift directions for summer months at this latitude agree with the results of YERG but not with those of TOMAN. During the winter months the drift directions observed here agree with those of YERG. It must be noted that both YERG and TOMAN have not given the exact hours when the observations were taken and as such, no detailed comparison was possible.

2.4. *Diurnal and seasonal changes in F2 drifts.*

We shall first take up the analysis of results obtained from diurnal study of *F2* drifts which involve 24 hr observations taken for 2–4 days in each month. Following the method of BRIGGS and SPENCER (1954) the N–S and E–W components for each observation are plotted against the hour of observation. Points for the observations of all the days in each month are put in one graph and mean curves are drawn by eye to represent the general trend of these points. From these curves average hourly values of N–S and E–W components of velocity are obtained and these smoothed values are then used for harmonic analysis. The results of the harmonic analysis for obtaining the steady part and 24 hr solar variation in the N–S and E–W components of drifts obtained for the two years 1956 and 1957 are shown in Table 2. The last column in the table gives the phase-angle difference between the N–S and E–W 24 hr components. Besides the analysis of data for each individual month, observations for all days in each season are put in one plot and mean curves are also drawn by eye showing the general trend of variation of the points for each season. From these smoothed curves hourly values of the velocity components are obtained and these values are subjected to harmonic analysis for obtaining not only the steady and 24 hr solar components, but also the 12 hr solar component. The results of the harmonic analysis of these curves for the years 1956 and 1957 are presented in Table 3.

To illustrate the general trend of variation the smoothed curves thus obtained for all the four seasons during 1957 only are reproduced in Fig. 4.

From a careful study of the results of the analysis of *F2*-drift data taken over the last two years we note the following chief points of interest for the *F2*-region drifts:

(1) The diurnal variation curves in Fig. 4 and other similar curves obtained for individual months show that for the *F2*-region drifts the variations have a predominant period of 24 hr rather than 12 hr.

(2) Except for minor variations, the general shape of the N–S and E–W diurnal variation curves for all the 3 months in any particular season is found to be the same. Very rarely we find deviations such as in the case of curves for the month of May 1956 for which the amplitude of the 24 hr period of the N–S component is low and the steady component in the E–W diurnal curve is high.

(3) In all the months and seasons the steady component of drift is very small. This is particularly so for the case of the N–S component. The steady part of the drift in both the N–S and E–W components is higher in the winter season.

(4) From Tables 2 and 3 it may be seen that the amplitudes of drifts having 24 hr period in the N–S component lie in the range of 65 m/sec to 75 m/sec while the corresponding values in the E–W component lies in the range of 80 m/sec to 90 m/sec. Thus in all the curves for the individual months and the seasons, the

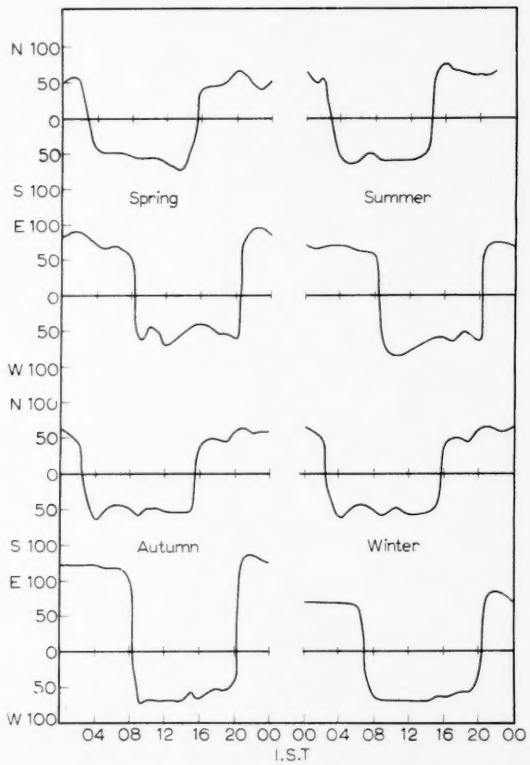


Fig. 4. Diurnal variation of N–S and E–W components of *F2*-region drifts in all the four seasons for the year 1957.

amplitude of the 24 hr solar cycle in the N–S component is always less than that in the E–W component.

(5) The phase angle of the N–S component is found to be invariably higher than that for E–W component. This sense of phase-difference indicates a drift direction which is changing continuously in a clockwise sense with time.

(6) The phase-difference between the N–S and E–W components shows slight variations from month to month but the major part of this variation is mainly statistical in nature owing to the uncertainty involved in drawing a curve through the scatter of points obtained for a season or for a month. In general it may be said that there is no significant change in the values of this phase difference from season to season, the values being in the range of from 70° to 80°. This phase-

difference is a consequence of the time lag of the E–W component behind the N–S component by about 5 hr which can be clearly seen from the curves in Fig. 4.

(7) The velocity vector of *F2*-drifts traces an ellipse in 24 hr thus going through maxima and minima twice a day. This feature can be easily seen from Fig. 5 which is a polar plot of the observations taken for three quiet days in the month of June 1957. The observations are shown alongside the smooth curve of the ellipse drawn by combining the steady and 24 hr components obtained from the harmonic analysis of the N–S and E–W curves for the month. It can be seen that the actual observations follow the trend of the elliptic curve fairly closely. The ellipticity, however, is found to vary slightly from month to month and season to season. The time in IST is shown alongside the points calculated from harmonic analysis results. In general it may be said that the drift direction which is generally towards NE at midnight changes rather rapidly through SE to SW in the morning hours. It remains SW for a fairly large duration during daytime and then changes over rapidly through NW in the late evening hours to NE during night-time.

Table 4. *F2*-Drift directions for 1956 and 1957

Season	0900	1200	1500
Spring	SW, SE*	SW	SW
Summer	SW	SW	SW, NW*
Autumn	SW	SW	SW, NW*
Winter	SW, SE*	SW	SW

VOL. 14 1959

(8) The component having 12 hr period in the N–S and E–W curves is not significant as the amplitudes generally lie in the range of from 3·0 to 16·0 m/sec excepting during winter, when it is larger and also as the phase differences do not show any consistent value.

(9) The main features of the diurnal variation curves and also the values for the amplitudes and phase angles obtained for the data taken in 1956 are found to repeat again during 1957. In general there is no significant variation in *F2*-drifts from 1956 to 1957.

F2-drift observations taken at 0900, 1200 and 1500 hours IST for a large number of days during the last two years were plotted as polar diagrams for each season and hour and the main drift directions obtained from these polar diagrams are shown in Table 4. Other directions of secondary importance are noted with asterisk.

It will be noticed that the noontime drift is SW for all seasons and this finding agrees with the conclusions from diurnal study as seen from Fig. 4. It is interesting to note that in the spring and winter seasons SE drift direction is also prominent at 0900 hours. This is explained as due to the transition of the E–W component being around 0900 hours. If this transition-time happens to be much earlier, as in the case of other two seasons, observations indicating SE direction at 0900 hours are not obtained. In a similar manner, a secondary direction towards NW is noticed for the 1500 hours observations during summer and autumn and this is

also explained as due to transition of N–S component occurring earlier in these seasons so as to be around 1500 hours.

Though fairly extensive *F2*-drift data is available for Cambridge as reported by Briggs and Spencer, results of harmonic analysis of this data are not available. It will be interesting to compare the results of this investigation with the main conclusions of Briggs and Spencer. Briggs and Spencer have reported that *F2*-drifts show a predominant diurnal variation of 24 hr period, a feature which is in excellent agreement with our findings. In all seasons at Cambridge the E–W component is towards west during night and east during day and this observation is exactly opposite to the behaviour of drifts at our station, where we find the

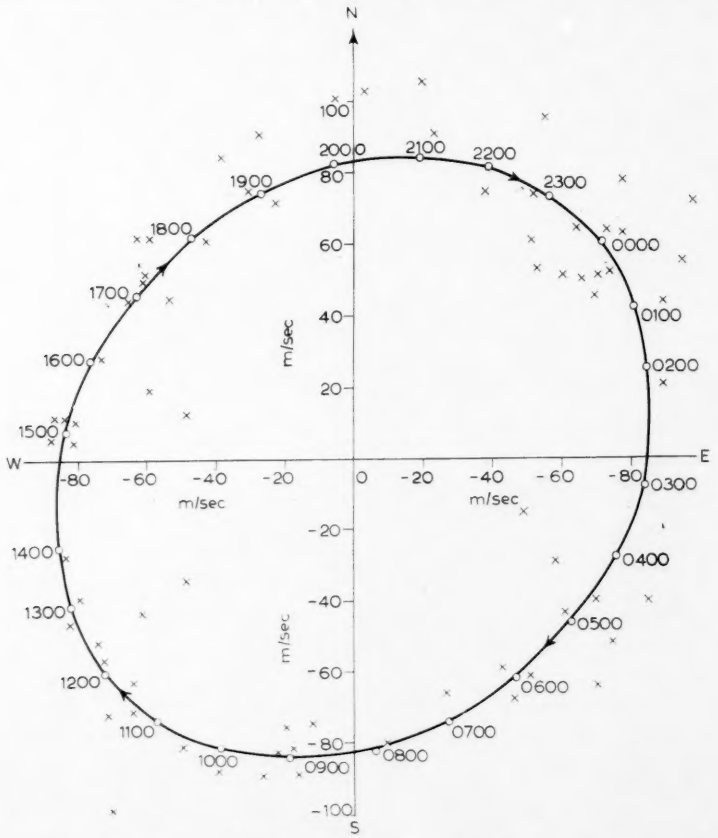


Fig. 5. Polar plot of *F2*-region drift observations for June 1957 alongside the elliptic trace obtained from the result of harmonic analysis.

E–W component directed mostly towards west during the day and east during night. Radio astronomy methods sometimes reveal a reversal of E–W component near midnight so that the velocity is towards east after midnight. It will be evident from the N–S and E–W curves for all the seasons given in Fig. 4, that such a reversal from west to east takes place at our latitude a few hours before midnight which varies with the season. Cambridge results show that the N–S

component is very small during night but no such large difference between day and night values is noticed in our observations. The N–S component at Cambridge shows during the day a significant seasonal variation, being towards south in winter and north during all the other seasons. It is interesting to note that we have not observed any such seasonal change in the N–S component. BRIGGS and PHILLIPS (1952) reported that the E–W component is usually greater than the N–S component, an observation which is in agreement with similar features noticed from our data for the last 2 years.

It is generally agreed that both in *E*- and *F2*-regions the variable part of the drift velocity vector rotates clockwise in northern hemisphere and anticlockwise in the southern hemisphere. Our observation is quite in agreement with this general conclusion. However harmonic analysis of *F2*-drift components has shown that the drift vector with 24 hr period, traces an ellipse giving two maxima and minima for the drift speed in a day as shown in Fig. 5. The curve shows maxima in drift speed at about midnight and noon. It is interesting to note that such maxima in drift speeds at about noon and midnight have also been observed by a simple plot of drift speed vs. time for all months and seasons as shown in Fig. 3.

The only other low-latitude station for which *F2*-drift measurements are available is Singapore (1°19′N, 103°49′E and mag. lat. 10°S) where OSBORNE (1955) has found a predominent semi-diurnal drift of amplitudes 40 m/sec from a limited number of observations taken from Sept. 1953 to April 1954. His results are not in general agreement with our findings. However, very recently PURSLOW (1958) has studied *F2*-drifts at Singapore from measurements taken over a period of 3 years and reported that the N–S and E–W components show a predominant 24 hr period. He found the E–W component of *F2*-drifts to be directed consistently towards east during night with a maximum amplitude of 90 m/sec and less consistently towards the west during the day with a maximum amplitude of about 30 m/sec. The N–S component is found to be much smaller than the E–W component at all times with consistent southerly direction during daytime and less consistent northerly direction during night-time. The general directions of the *F2*-drift components observed by PURSLOW are in good agreement with ours although they are exactly in anti-phase to those determined at higher latitudes. Though we find that the night-time values of the E–W component are in general slightly higher than the E–W components during daytime, the large divergence between the night-time and daytime E–W drifts as reported by PURSLOW has never been observed by us. It appears that the magnitude of southerly drifts observed during daytime at Singapore is much higher than the northerly drifts during night-time, while at our station we do not find any significant differences in the amplitudes of the N–S component from night to day. The amplitudes of the 24 hr components as reported by PURSLOW from the harmonic analysis are found to be much less than the values obtained by us. As the phase difference between N–S and E–W 24 hr components appears to be about 1 hr, from the harmonic analysis data given by PURSLOW, it will be evident that the 24 hr drift vector at Singapore rotates in a clockwise sense tracing an ellipse of high eccentricity. The eccentricity of the ellipse at this latitude is much less than this as the phase difference is of the order of 5 hr.

VOL. 14 1959

The drift movements at this latitude are exactly in opposite direction to those observed at higher latitudes, thus confirming the prediction of Martyn (1955) that there should be a phase reversal of *F2*-drifts at a latitude of 35°. It is interesting to note that the amplitude of the 24 hr E–W component is much higher than the corresponding values obtained for higher latitudes. Using separate expressions for the eastward drift velocity around the magnetic equator and at high latitudes, Martyn had shown that the maximum eastward drift by day at Cambridge (52°13′N, 00°6′E and mag. lat. 54°44′) should be about 35 m/sec and the maximum westward drift by day near the magnetic equator should reach a value of 200 m/sec. The maximum westward drift of 80 m/sec to 90 m/sec observed at this latitude during daytime falls in between these two values, confirming in a way the predictions of Martyn concerning the variation of maximum E–W component of drift with latitude.

Acknowledgement—We are indebted to the Council of Scientific and Industrial Research of India for financial support for these investigations.

References

Briggs B. H. and Phillips G. J.	1952	*Proc. Conf. on Ionospheric Physics* (*July*, 1950). Geophys. Res. Paper II, p. 199.
Briggs B. H. and Spencer M.	1954	*Rep. Progr. Phys.* **17,** 245.
Martyn D. F.	1955	*Physics of the Ionosphere* pp. 163 and 260. Physical Society, London.
Mitra S. N.	1949	*Proc. Instn. Elect. Engrs.* Pt. III, **96,** 441.
Osborne B. W.	1955	*J. Atmosph. Terr. Phys.* **6,** 117.
	1951	*J. Atmosph. Terr. Phys.* **2,** 66.
Purslow B. W.	1958	*Nature, Lond.* **181,** 35.
Rao B. R., Rao M. S. and Murty D. S.	1956	*J. Sci. Industr. Res.* A **15,** 75.
Toman K.	1955	*J. Geophys. Res.* **60,** 57.
Yerg D. C.	1955	*J. Geophys. Res.* **60,** 173.
Wells H. W. and Stanton H. E.	1938	*Terr. Magn. Atmos. Elect.* **43,** 169.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 107 to 110. Pergamon Press Ltd. Printed in Northern Ireland

Determination of the angle of arrival of auroral echoes

L. HARANG and J. TRÖIM
Norwegian Defence Research Establishment, Division of Telecommunication,
Kjeller, Lillestrøm, Norway.

(*Received* 21 *May* 1958)

Abstract—The angle of arrival, θ, of auroral echoes is measured by an interference method. The variation of θ with echo range R is demonstrated. θ varies from 15° to 6·5° when the range increases from 400 to 730 km. The height of the reflection area must lie at 100–120 km.

VOL.
14
1959

INTRODUCTION

ALL observations of auroral echoes indicate that they are coming in at low angles with an angle of arrival of less than 15–20° (CURRIE *et al.*, 1953; LITTLE *et al.*, 1956). In the frequency range 40–80 Mc/s in which most observations of auroral echoes are made, it is not practically possible to construct an aerial with directivity sufficient to avoid the presence of ground reflections at these low angles of arrival. The estimates of angle of arrival have, therefore, hitherto been based on indirect evidence.

Methods have, however, been in use for measuring low angles of arrivals of other phenomena. The method is based on measuring the amplitudes of signals simultaneously in two aerials placed vertically above each other at certain heights and separations. From the amplitude ratio between the signals in the two aerials the phase difference between the direct and the ground reflected wave can be computed, from which the angle of arrival can be evaluated. The method has been used for measuring low-angle arrival of CW signals reflected from the E- and F-layers (WILKINS and MINNIS, 1956), and of transient signals the angle of arrival of meteoric echoes have been studied (ASPINALL *et al.*, 1951). A similar method has now been used for measuring the angle of arrival of auroral echoes, but due to the irregular and rapidly varying character of these echoes both in structure and range the method had to be modified.

METHOD OF MEASUREMENT

Conventional echo transmitter and receiver working on 40 Mc/s with a peak power of 20 kW were used. The same aerial was used both for transmission and reception, the feeder being supplied with a R–T coupling. The electric vector was horizontal in transmission and reception. If we have a perfectly reflecting horizontal plane in front of the aerial the reflected ray will get a phase change at reflection of 180°. If the direct and reflected rays are received (with horizontal polarization) at a point lying in height h above the ground the resultant e.m.f. is given by $e = 2E \sin\left(\frac{2\pi h \sin\theta}{\lambda}\right) \cos\omega t$, where θ is the angle of arrival. If we have a constant angle of arrival and are moving the receiving point upwards, we will record a system of interference fringes with spacings determined by the angle θ. The experiment with this modification represents an analogy to LLOYD'S fixed-

mirror interference experiment. Fig. 1 illustrates the experimental arrangement. The frequency used was 40 Mc/s. Some tests have also been made on 80 Mc/s. All observations were made at Kjeller ($\varphi = 60°$N) and the aerial was directed towards north with approximately 10° to west.

In the first series of experiments a fixed yagi was used as transmitting aerial and only the receiving aerial was moved up and down. The echoes were isolated

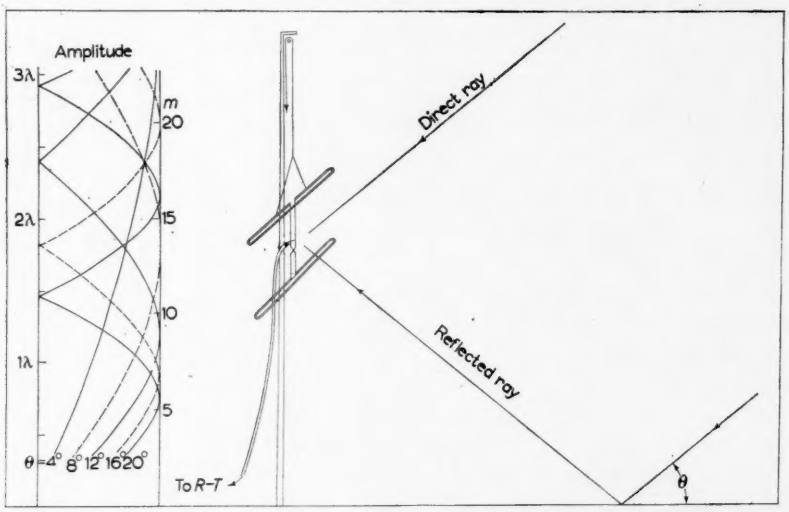


Fig. 1. The receiving + transmitting aerial could be moved continuously up and down. The amplitudes of the auroral echoes will vary according to the curves to the left, depending on the angle of arrival θ.

through a gate and the mean amplitudes were recorded through a peak voltmeter. It turned out, however, that this procedure gave inconsistent results due to the lobe effect of the fixed transmitting aerial. In the later experiments the moving aerial was used both for transmitting and receiving. The width of the echoes usually lies in the interval 50–150 km, which may represent different ranges and angles of arrival. It was therefore necessary to use a very narrow gate and to isolate only a small portion of the auroral echo pattern. The range corresponding to the position of the gate was simultaneously recorded in the echo pattern. The amplitude of the section of the auroral echo pattern isolated by the gate was recorded through a second cathode-ray tube.

The aerial, two folded dipoles mounted broadside in $\lambda/2$ distance, was used both for transmission and reception working through a R–T switch. The aerial could be moved up and down with constant speed along the mast. In front of the aerial a horizontal field was used as the reflector. For very low angle radiation ($h = 2\lambda$ and $\theta \sim 6°$) this requires a plane, horizontal field in front of the aerial extending about 350 m to the north, for covering the first Fresnel zone. The field in front of the aerial did not quite fulfil this condition, but no serious error should be introduced. It is, however, obvious that values of θ less than 6° should be viewed with caution.

VOL. 14 1959

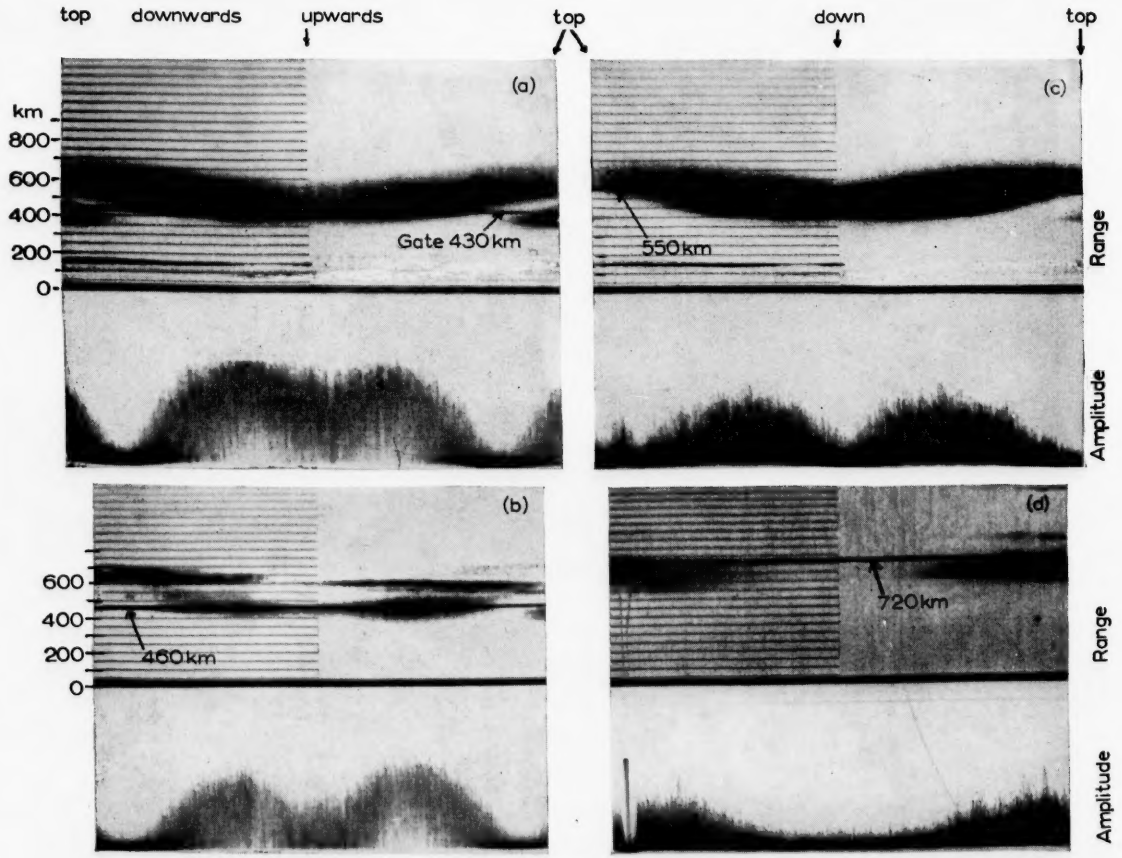


Fig. 2. Records for determination of angle of arrival of auroral echoes on 40 Mc/s at Kjeller ($\varphi = 60°$N). The upper part of each record gives the range. The lower part gives the amplitude at a given range within the echo pattern, indicated by the "gate-line". During each record the aerial has been moved uniformly from the top, down and up along the mast.

VOL
14
195

In Fig. 2 four records are shown giving ranges and amplitudes of the auroral echoes during one period of movement of the aerial up and down. The records are taken with two cathode-ray tubes, the upper part of the records gives the range and the lower part the amplitudes in the echo pattern at a range indicated by the horizontal gate-line. The nearest echo (*a*) at 430 km range exhibits both a maximum and a minimum value in the amplitude variation from which an angle of arrival of 12·5° can be deduced. The most distant echo (*d*) at a range of 720 km exhibits only a flat maximum near the top of the aerial from which an angle of arrival of 6·5° can be deduced.

Discussion

A series of determinations of θ as a function of R is given in Fig. 3. Although there is some scattering of the individual points the figure indicates that the height of reflection above the earth's surface must lie at a height of 100–110 km for echo

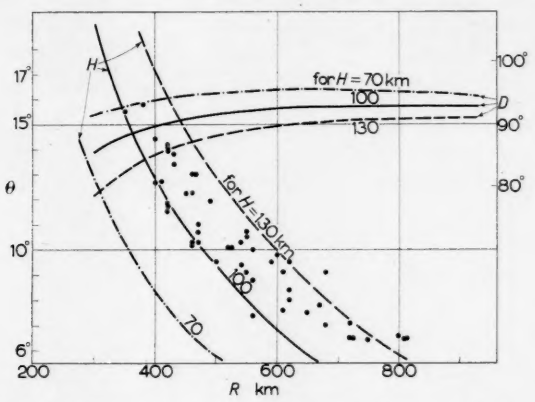


Fig. 3. Angle of arrival, θ, as a function of range, R. If it is assumed that the reflections take place alternatively at heights H = 70, 100 or 130 km, the angle of arrival should lie on the curves indicated. The three horizontal curves show the variation of D, which represents the angle between the geomagnetic lines of force and the direction of arrival.

ranges 400–500 km. For ranges 600–700 km the heights seem to increase slightly, to 110–120 km. The variation of θ with R seems to indicate that for the range interval considered there has been no appreciable bending of the ray during the passage of the lowest part of the E- or E_s-region. At high-latitude stations close to the auroral zone where ranges down to 200 km appear, this effect may have to be considered.

Further the angle D has been calculated, representing the angle between the direction of the geomagnetic lines of force and the direction of arrival (Chapman, 1952)* for three heights of reflection, 70, 100 and 130 km. For the echoes in ranges 400–500 km D attains a value very near to 90° for an assumed reflection-height of 110 km. For the more distant echoes in 500–700 km range the reflection-heights must lie in somewhat greater heights to fulfil the condition $D = 90°$. The variation of θ with range seems to confirm this point.

* As values for the dip, mean values for Rude Skov ($I = 69·0°$) and Tromsö ($I = 77·5°$) have been used, and linear interpolations have been made to obtain values at the points of reflection.

The observations here described were made at Kjeller ($\varphi = 60°$N) which represents a station well south of the auroral zone. It is planned to extend these measurements to the arctic station Tromsö ($\varphi = 70°$N) lying close to the aurora zone and where auroral echoes seem to exhibit a much greater variation in range and complexity.

Acknowledgement—The authors wish to express their thanks to R. BRATTENSBORG for his assistance during the observations.

REFERENCES

ASPINAL A., CLEGG J. A. and HAWKINS G. S.	1951	*Phil. Mag.* **42,** 504.
CHAPMAN S.	1952	*J. Atmosph. Terr. Phys.* **3,** 1.
CURRIE B. W., FORSYTH P. A. and VAWTER F. E.	1953	*J. Geophys. Res.* **58,** 179.
LITTLE C. G., RAYTON W. M. and ROOF R. B.	1956	*Proc. Inst. Radio Engrs., N.Y.* **44,** 992.
WILKINS A. F. and MINNIS C. M.	1956	*Wireless Engr.* **33,** 47.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 111 to 122. Pergamon Press Ltd. Printed in Northern Ireland

On the aurora of 19 August 1950, photographed from Greece

W. N. ABBOTT* and S. CHAPMAN†

(*Received* 30 *May* 1958)

Abstract—Further details are given concerning an auroral photograph already reproduced and discussed (ABBOTT, 1951). The association between the occurrence of auroræ in low latitudes and the 3 hr geomagnetic disturbance index K_p is considered.

The aurora was reported from places over a wide area (Europe, the Atlantic and North America); its reported appearances ranged from 18 August at 2130 hours to 19 August, 0530 hours UT (at Nantucket); and from New Zealand it was seen after sunset, from about 0730 UT on 19 August.

VOL. 14 1959

1. INTRODUCTION

A SKY photograph of exceptional interest, taken by one of us, was reproduced and briefly discussed in an earlier note in this journal (ABBOTT, 1951). The present note amplifies and partly supersedes that discussion.

The plate, showing part of the eastern sky not long before dawn, was taken in the course of a study of the autumn reappearance of the zodiacal light; the moon was 3 days old, hence it was further below the horizon than the sun. Some details are as follows:

Place: Spetsai College astronomical station, Spetsai Island, Greece. 37°13′N, 23°17′E geographic (gg); gm (geomagnetic) latitude 35°57′N.

Time: 19 August 1950; exposure 25 min, centred at 0207 hours UT (0407 hours local standard time). The exposure ended shortly before astronomical twilight began.

Camera: F : 4·5 Goerz lens; attached to a telescope with equatorial drive.

Plate: Kodak commercial Super-XX Panchromatic; sensitivity uniform in the red and green, increasing in the blue and violet.

Star-field: Gemini (Uranus also appears on the plate).

When the exposure was stopped a glow, not previously seen, was casually noticed in the part of the sky to which the camera was directed. This was mistaken for the zodiacal light, until the plate, after development, revealed the complex pattern of the light. Fig. 2 is a (negative) pen drawing of the original plate; black represents brightness. Unlike the positive photographic copy given in the earlier note, it is drawn so as to show the location of the luminosity in azimuth and elevation, in which its ranges were from about 75° to 87°E of N and from about 7° to 20° elevation. At Spetsai the geomagnetic East is in the geographic direction 75·6°E of N (as marked on Fig. 2). Thus the luminosity lay in a geomagnetic latitude south of Spetsai, that is, below 36°N.

* University of Athens, temporarily at the Geophysical Institute, College, Alaska.

† Geophysical Institute, College, Alaska; High Altitude Observatory, Boulder, Colorado (there working with the support of the National Bureau of Standards and the Air Force Cambridge Research Center.)

2. Auroral Observations, 19 August 1950

When the plate was examined and the unexpected image of sky luminosity was found on it, the image was not immediately recognized as auroral. However, news was later received that a great aurora had been seen on that date. In fact, 1 hr earlier a brilliant aurora had been seen from the British liner *Queen Elizabeth*, bound from Southampton to New York. The ship was then at 42°41′N, 52°41′W (53°N gm). From 0059 hours UT continuous waves of brilliant light, resembling powerful searchlight beams, were seen travelling across the sky from N to S, dominating the entire sky, and changing direction from time to time. At 0115 hours UT waves of light streaked vertically upwards from the entire horizon, to culminate in a white glow at the zenith. Towards the southern side of the apex a dull reddish glow was noticed, fading after a few brief moments. (Report by First Officer G. E. Smith; published in U.S. Hydrographic Bulletin No. 3187 of 7 October 1950, and *J. Geophys. Res.* **55,** 498, 1950.)

The southern border of this "Atlantic" aurora apparently lay below the horizon of the ship. If the aurora extended down to 100 km its minimum geomagnetic latitude must have been below 46°, because an object at 100 km height visible (from sea-level) on the horizon lies at a distance of more than 800 km or 7° of latitude from the observer. Low latitude aurorae may not extend so far down as 100 km. If this aurora was wholly above 100 km, it must have been at a gm latitude less than 46°.

This "Atlantic" aurora is notable as showing a corona visible from 53° gm latitude. (One may recall that during the outstanding aurora of 4 February 1872, a corona was seen from London, 54° gm latitude.)

It is interesting to note that 19 August 1950 was the first of the two nights on which Meinel (1951) obtained his important spectrograms showing Doppler displacements of the *H*-line. His work began on the evening of 18 August local time (at the Yerkes Observatory: 55° gm). Auroral spectrography could not have begun at the Yerkes Observatory much, if any, before 8 p.m. local time, 18 August, which was 19 August 0200 hours UT, the time of the auroral photograph discussed in this note. On that night the aurora apparently did not come overhead at Yerkes; on the next night it passed the zenith and went south at least to the magnetic zenith (20° south of the true zenith), and a corona was observed. Other auroral observations on the night of 18–19 August have kindly been brought to our notice as follows by Miss C. M. Botley and Mr. D. S. Kimball:

(1) Ocean Weather Ship *Weather Explorer*. 58°15′N, 15°25′W (63° gm); 19 August, 0001 hours UT. Marked aurora consisting of diffuse green coloration extending to the zenith with some bands south of the zenith. The duration of the display is not given, but presumably the aurora lasted for some time, as the observer reports that it remained fairly constant in pattern. Slight aurora was also reported at 0300 hours UT.

(2) *S.S. Essex Trader*. 57·5°N, 18·0°W (62° gm); 19 August, 0150 hours–0400 hours UT. Reported as having observed rays (HA, RA, RR) extending from east to west, presumably passing overhead.

(3) *Whickam, England*. 54·6°N, 01·4°W (54° gm); 18 August–19 August; red glow.

(4) *Petersfield, England*. 51·0°N, 00·9°W (54° gm); 18 August, 19 August; red glow.

(5) *Coulee Dam, Wash., U.S.A.* 48·0°N, 114°W. (54° gm); 18 August, 2130 hours to 19 August, 0250 hours UT. Ray formations to the zenith reported but no corona.

(6) *Battle Creek, Mich., U.S.A.* 42·3°N, 83°W (53 gm); 19 August, 0125 hours–0437 hours UT. Reported types G, DS, R, D; coronae at 0330 hours, 0352 hours, 0357 hours, 0411 hours UT.

(7) *Moline, Ill., U.S.A.* 41·3°N, 90·3°W (50° gm) 19 August, 0315 hours–0345 hours UT. Reported faintly luminous cloud to north.

(8) *Nantucket, Mass., U.S.A.* 41·3°N, 70°W (52° gm); 19 August, 0155 hours–0530 hours UT when clouds interfered. Types recorded, RA, R, D, PA, PS; corona at 0200 hours, 0235 hours, 0237 hours, 0308 hours UT. Bright.

(9) *Polk, Ohio, U.S.A.* 41·0°N, 82°W (52 gm); 19 August, 0300 hours UT. Aurora was recorded overhead.

(10) *Middleton, Ohio, U.S.A.* 39·5°N, 70°W (50° gm); 19 August, 0150 hours–0325 hours UT. Types recorded, G, HA, RA, R, PS, F. Corona at 0305 hours UT. Bright.

Fig. 1 gives the locations and the places whence these auroral observations in the northern hemisphere were made; it is centred on the geomagnetic pole (at 78·5°N, 69°W, geographic).

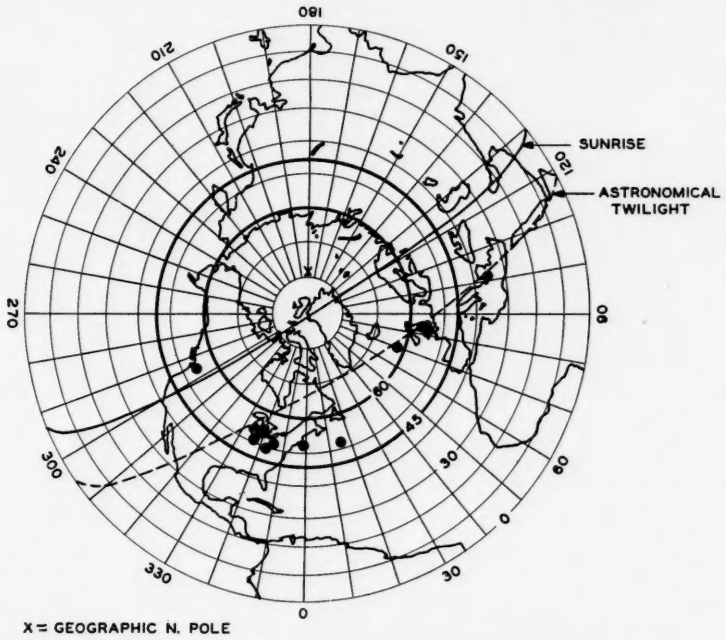


Fig. 1. The northern hemisphere centred on the northern geomagnetic pole (at 78·5°N, 69°W geographic). Geomagnetic meridians and parallels of latitude are shown. The gm parallels 45°N, 60°N are the boundaries of the northern subauroral belt; the gm parallel 60°N contains the northern auroral region. The northern half of the minauroral belt is shown as well as the demarcation line between sunlight and darkness (sunrise line) and the line of astronomical twilight (dashed line), at 0207 hours UT, 19 August 1950. Dots (●) mark the location of observers mentioned in the text.

Geomagnetic meridians and parallels of geomagnetic latitude are shown. The parallels 45°N, 60°N are the boundaries of the northern subauroral belt; the parallel 60°N contains the northern auroral region. The northern half of the minauroral belt is shown. The aurora photographed from Spetsai was observed from, and lay in, the minauroral belt. The other observations of auroras in the northern hemisphere which appeared at about the same time, i.e. between 0000 hours and 0400 hours, were made from the northern subauroral belt, with the exception of that from the *Essex Trader* which was made from within the auroral zone.

The geomagnetic planetary 3 hr index K_p for the period 0000 to 0300 hours

VOL. 14 1959

on 20 August 1950 was 6−; the values for the 24 hr from noon (UT) 18 August to noon 20 August 1950, were as follows (BARTELS and VELDKAMP, 1951):

Noon					Midnight				Noon
K_p		18 August				19 August			
	2_0	3+	4_0	5−	6−	4+	6−	6+	
K_p		19 August				20 August			
	8_0	8+	8−	8_0	8+	8_0	8−	8−	

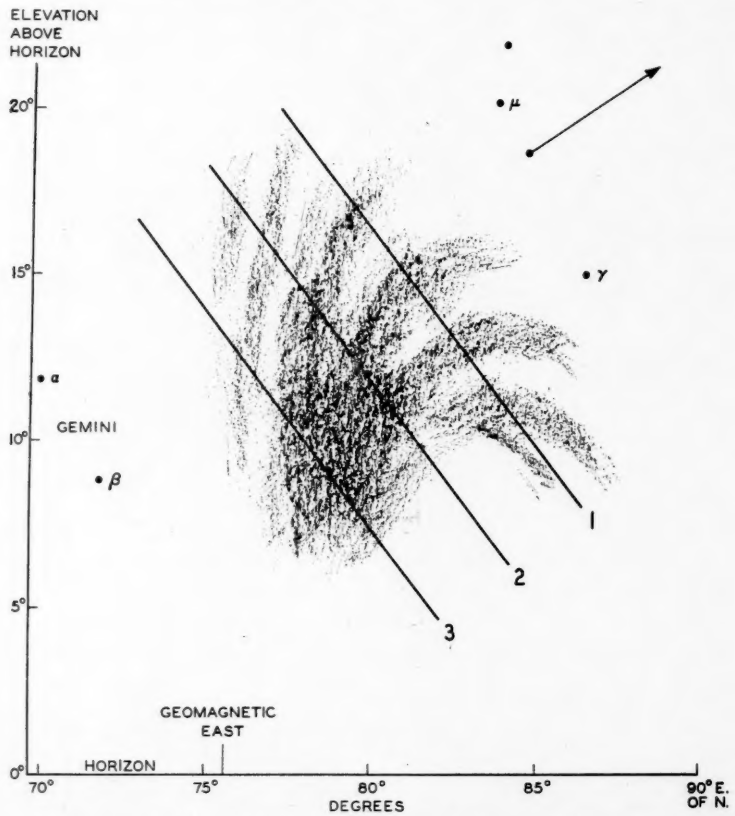


Fig. 2. The aurora of 19 August, 1950, photographed in Greece (*J. Atmosph. Terr. Phys.* **1**, 343, 1951). Microphotometer scannings were made along the lines 1, 2, 3. Ordinates are elevations, abscissae are azimuths. The star-field is located for the end of the exposure. The arrow shows the length of the 25 min star-trails had the camera been motionless.

The high K_p value at the time of the Spetsai photograph, and the near coincidence in time between it and the "Atlantic" aurora, leave no room for doubt that the image on the Spetsai plate represented an aurora. Hence the photograph is of outstanding interest; it appears to be the first objective picture of a really low-latitude aurora (geomagnetic latitude below 36°N) (CHAPMAN, 1957a, b). The following night, 19–20 August, a period of higher K_p values, was one of notable

aurorae (see *L'Astronomie*, 1950 and 1951); the first appearance of this aurora as reported to the U.K. Meteorological Office was from *S.S. Fort Spokane*, 39° 04′S, 146° 40′E, whence a good display of aurora australis was seen to the south at 1300 hours UT.

The aurora of 18–19 August was seen from New Zealand (as kindly reported to us by I. L. THOMSEN), beginning about sunset there at 19 August, 0747 hours (Campbell Island) and 0800 hours UT (New Zealand).

3. The Auroral Surface Brightness

The low gm latitude of the auroral appearance photographed at Spetsai justifies an attempt to estimate its surface brightness. A close determination is precluded because of the uncertain duration of its appearance during the photographic exposure. Fig. 2 shows by an arrow the extent and direction of motion of the image of a sky point in a fixed direction relative to the moving camera and the plate. The motion is large relative to the fine detail of the auroral image. Hence the aurora must have been registered for only a brief final portion of the exposure, certainly less than 5 min and probably only 1 or 2 min.

The following method of estimating the surface brightness of the aurora was suggested by Dr. C. T. ELVEY. Microphotometer tracings of the image intensity on the original negative were made along the lines 1, 2, 3 (shown on Fig. 2). The tracings are reproduced on Fig. 3, in which the abscissae are degrees of arc on the plate. The plate shows a background blackening attributable to the airglow, and this is assumed to have been of normal intensity; its intensity level, and the base-level of the clear plate (the zero of intensity is not very distinctly marked on the plate) are plausibly surmised as in Fig. 3. The mean surface brightness of the aurora, E_a, is consequently inferred to be 1·4 times the background airglow surface brightness, E_s. The average value of E_s at the zenith is about 250 R (R signifies rayleigh; 1 R denotes 4×10^6 quanta/cm^2 sec steradian) (HUNTEN, ROACH and CHAMBERLAIN, 1956; ROACH, 1957). At the elevation of the aurora in Fig. 2, namely about 10°, E_s is about twice as great, or about 500 R. If the auroral exposure was only 5 min of the full 25 min, E_a may be reckoned to be $25/5 \times 1{\cdot}4 \times 500$ R or 3·5 kR (kilorayleighs). If, as is thought more likely, the aurora was present during the exposure for only 1 min, then E_a is 5 times greater, namely 17·5 kR. Hence the aurora may be classed as moderately bright (ROACH, 1957, p. 115).

4. The Locations of the Aurora

Figs. 4 and 5 indicate alternative possible locations of the aurora photographed from Spetsai, in plan and in elevation (the latter projected on the vertical plane through Spetsai and the aurora). Fig. 4 shows also for the epoch at the end of the exposure the line of commencement of astronomical twilight, and the sunrise lines at ground level and at heights of 100 and 450 km. If the lower limit of the aurora was at 100 km, the aurora was about 500 km distant, and wholly in the shadow. If its distance was 1000 km, its lower border was at about 200 km height. Positions intermediate between those shown are, of course, also possible.

Low-latitude aurorae are generally thought to be at higher levels than the

VOL. 14 1959

usual high-latitude aurorae; if so, this was a sunlit aurora, probably located somewhere south of the Black Sea, over Turkey. Owing to dawn brightening, it may not have been noticeable from the regions over which it lay. But it should

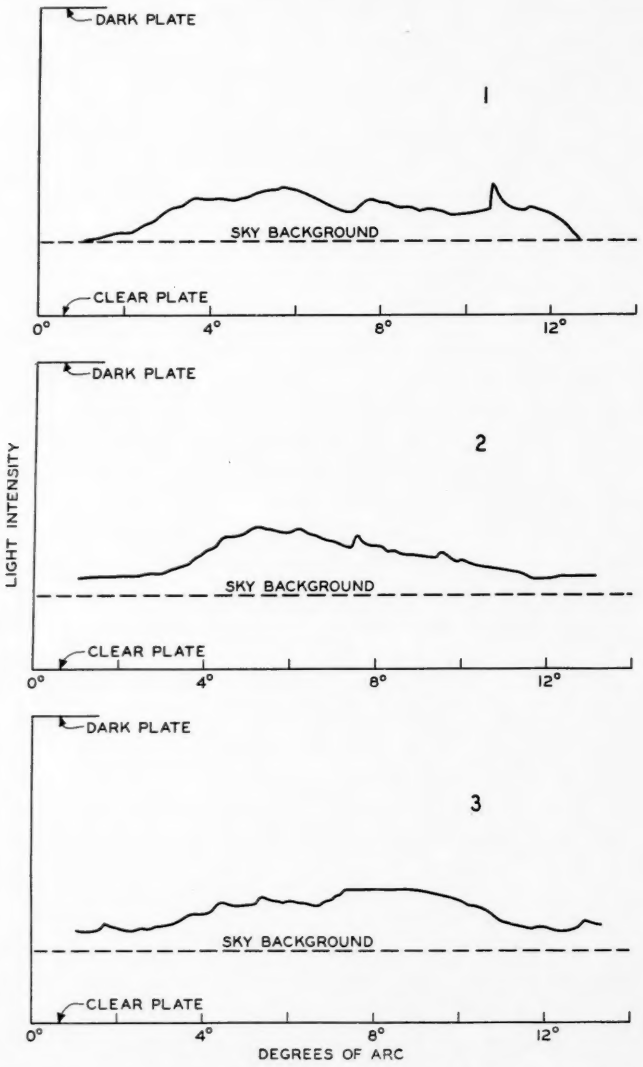


Fig. 3. Microphotometer tracings along the lines 1, 2, 3 of Fig. 1. The ordinates I are intensities and the abscissae are degrees of arc along the lines of scanning.

have been visible from much of Greece (including Crete), and Bulgaria, Rumania and western Turkey. We know of no record of the aurora photographed from Spetsai other than the photograph here discussed. However, aurorae were seen at about the same time from many other places.

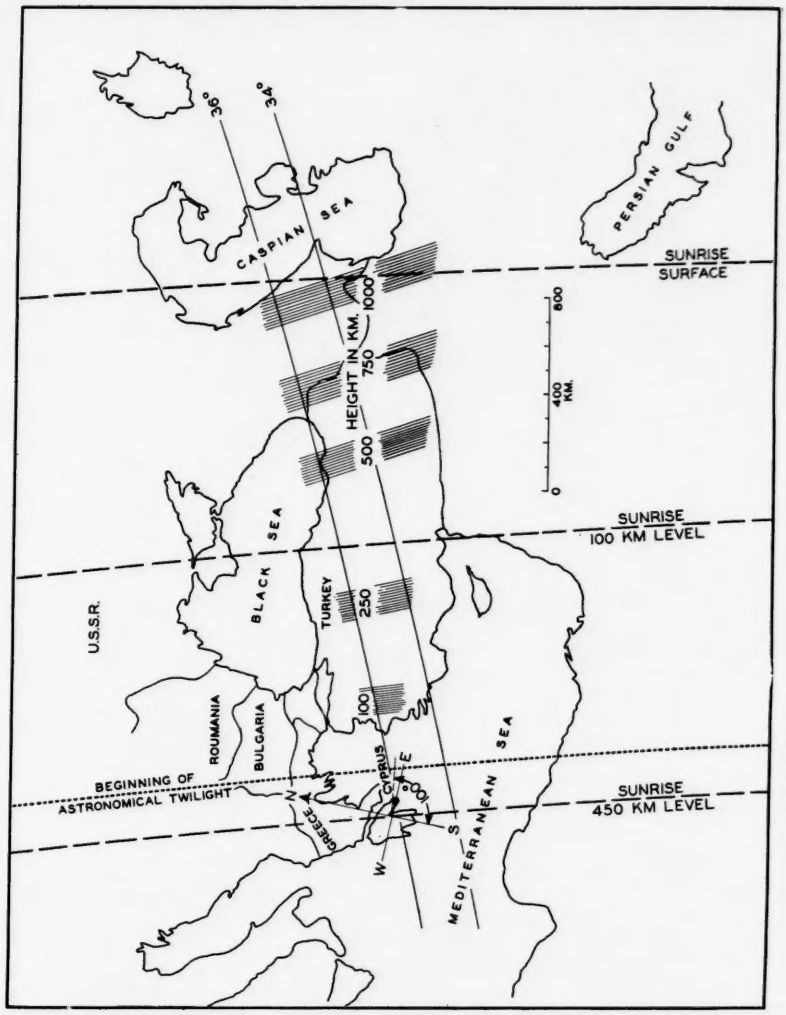


Fig. 4. Possible locations in space of the aurora photographed from Spetsai on 19 August 1950, for a number of assumed altitudes.

5. The Form of the Spetsai Aurora

The auroral image has a central region bordered to the left (the north) by a separate narrow band or streamer, and fainter streamers still further north; to the right (southward) there are fanlike streamers with intervening dark spaces. From a single photograph it is impossible to infer their true form. The streamers on the left cannot have lain along the lines of geomagnetic force.

6. Low-Latitude Aurorae and Geomagnetic Disturbance

There is undoubtedly a relation between the intensity of geomagnetic disturbance (D) and the occurrence of aurorae in unusually low latitudes—either

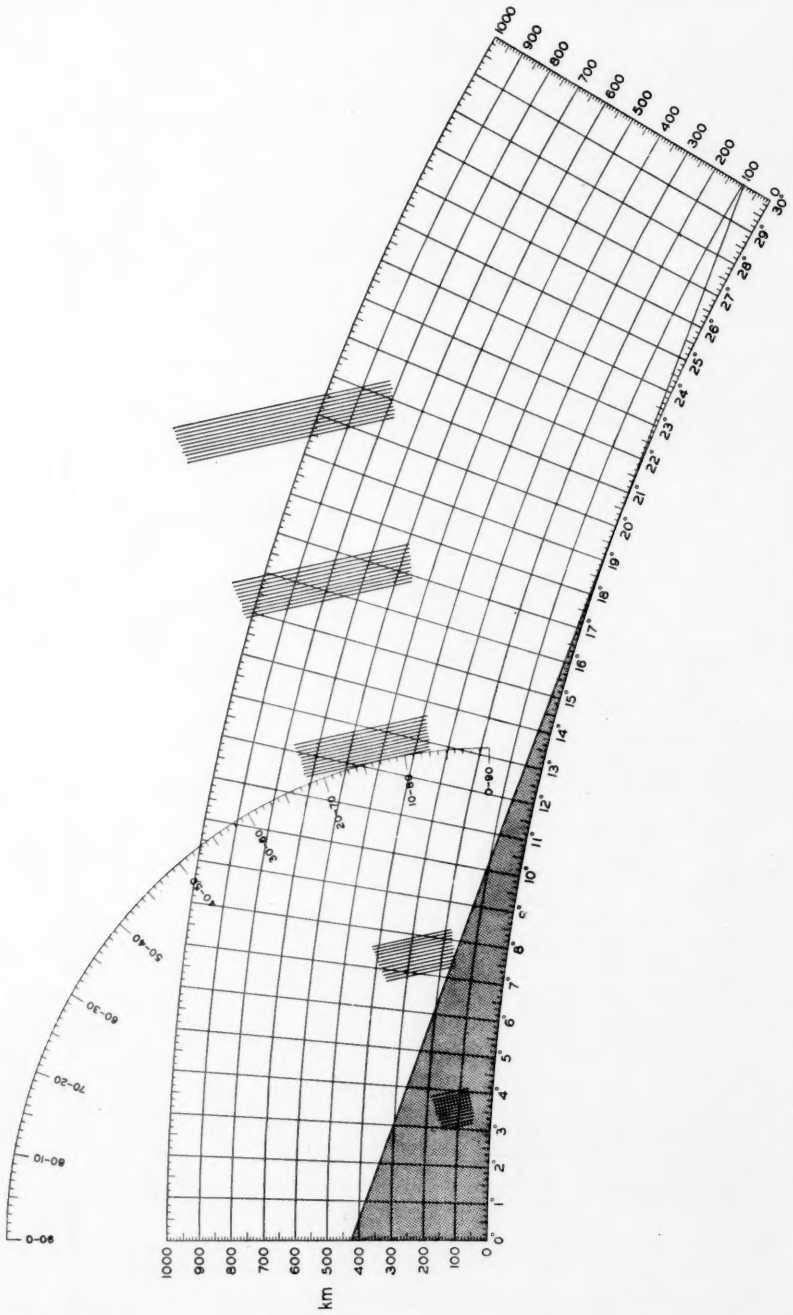


Fig. 5. Possible geographic locations of the aurora photographed from Spetsai on 19 August 1950, for the distances shown in Fig. 3.

subauroral gm latitudes, between 45° and 60°N or S, or minauroral (between gm latitudes 45° N and S). To facilitate discussion of this relationship, a recent paper (BARTELS and CHAPMAN, 1958) tabulates all the occasions during 21 years (1932/33 and 1937–1957) on which K_p, the planetary 3 hr index of geomagnetic disturbance, had values 5_0 or more. On the basis of this table we gave a diagram showing the greatest K_p value (not less than 5) in the four quarter-days centred at 0000, 0600, 1200, 1800 hours UT; this bears on the likelihood of low-latitude aurorae being visible in the four quadrants of the earth centred on the meridians of longitude 0, 90°W, 180°, 90°E. The phases of the mean moon are also indicated on this diagram, as moonlight affects the visibility of aurorae.

The greatest aurorae in the last century, of 28 Aug. to 2 Sept. 1859, 4 Feb. 1872, 25 Sept. 1909 and 13 May 1921, coincided with outstanding magnetic storms. But for lower degrees of D, as measured by K_p, the parallelism appears to be less close than may be generally supposed. The auroral data, however, are so scanty and incomplete that it is difficult to compare the incidence of the two phenomena.

During the International Geophysical Year the watch for the aurora is exceptionally widespread and careful. This has already provided considerably more records of aurorae visible from minauroral latitudes than would correspond to the well known diagram of frequency of auroral visibility given by FRITZ (1881) and revised by VESTINE (1944). But *visibility from* minauroral latitudes does not, generally, imply *location of the aurora in* such latitudes. The presence of the 19 August 1950 aurora in a geomagnetic latitude less than 36° is a decidedly exceptional event. Yet the associated K_p figure was only 6−. Although this denotes a considerable degree of disturbance D it is distinctly below what would justify an expectation of an aurora in so low a gm latitude. For the same 3 hr period on the following night, 19–20 August, K_p was much higher, 8−; this prompted one of us (S. C.) to enquire whether the Spetsai photograph could have been taken on the night *beginning* on 19 August, but the reply was definitely negative.

Values of K_p equal to or greater than 6− occur in about one 3 hr period out of 48, and are thus fairly common, though very irregularly distributed in time. Values 8+ or more (that is, 8+, 9−, 9) occurred for only 46 of the 3 hr periods in the 20 years 1937–56—that is, for only 1 out of 1270 periods, on the average. Yet we know of no observations of low-latitude aurorae on the night of 19–20 August, when K_p rose to 8+.

Other examples of considerable deviation from close parallelism between K_p and low-latitude aurorae may be cited: BARTELS and CHAPMAN (1958) quote the case of an aurora of 20 Oct. 1957 *located in* gm latitude 51°N and associated with $K_p = 4°$. BARBIER (1957) cites a case of $K_p = 5+$ without aurora visible from France. On 12–13 December 1950, between 2330 hours and 0130 hours UT, an interval associated with K_p values 4+, 5+, an aurora was seen from latitude 40°45′ N, 55°00′ W (gm latitude 52°); but as it rose only to an elevation of about 15°, it was further north than the 20 Oct. 1957 aurora.

The IGY auroral data will permit better examination of these questions. Their proper answer, however, will require the continuance of the worldwide

auroral watch for at least some years more. It may be that the appearance of auroras in low latitudes depends not only on the value of K_p at the time, but also partly on earlier values, and on the mode of variation of K_p.

7. The Need for Increased Auroral Vigilance in Low Latitudes

This aurora of 19 August 1950 may serve as a text to emphasize the need for increased vigilance and activity in auroral observation in low latitudes. One of us (W. N. A.) saw the aurora without realizing what it was. Accordingly, instead of taking further photographs of it, or watching it to observe its changes and duration, he immediately proceeded, as dawn was breaking, to develop the plate here discussed. Only later, on reading an account of the great aurora of that day, did he recognize what he had photographed. Thus he missed the rare opportunity of watching it (or further photographing it) during the remaining minutes of darkness. He was not on the alert for observing aurorae.

Not until many more people in low latitudes become aurora-conscious will the rare and intensely interesting low latitude aurorae be properly recorded. The Japanese National IGY Committee is to be congratulated on its installation of an all-sky camera at Tokyo (gm latitude 25°4′N), for use on the rare occasions when an aurora comes within its range. It is much to be hoped that this example will be followed elsewhere, and that meteorologists and astronomers will personally keep the auroral possibilities in mind and encourage others to do so. It is also important that preparation be made to determine the *spectrum* of unusual luminosities in the skies of low latitudes. In particular, it is desirable to ascertain whether the light includes components which require for their production an energy of excitation greater than that associated with the airglow.

Acknowledgements—We are indebted to D. C. Wilder of the Geophysical Institute, College, Alaska, for the preparation of the figures for this note.

References

Abbott W. N.	1951	*J. Atmosph. Terr. Phys.* **1,** 343.
Anonymous	1950	*J. Geophys. Res.* **55,** 498.
Barbier D.	1957	*C. R. Acad. Sci., Paris* **244,** 1945.
Bartels J. and Veldkamp J.	1951	*Bull. Int. Ass. Terr. Magn. Elect.* **12,** 105.
Bartels J. and Chapman S.	1958	*Abh. Akad. Wiss. Göttingen, Math. Phys. Kl.* In press.
Chapman S.	1957a	*Nature, Lond.* **179,** 7.
Chapman S.	1957b	*Ann. IGY* **4,** 37.
Hunten D. M., Roach F. E. and Chamberlain J. W.	1956	*J. Atmosph. Terr. Phys.* **8,** 345.
Roach F. E.	1957	*Ann. IGY* **4,** 115.
Meinel A. B.	1951	*Astrophys. J.* **113,** 50.
Vestine E. H.	1944	*Terr. Magn. Atmos. Elect.* **49,** 77.

Note Added In Proof

Several leaders in auroral observation have been consulted in the hope of obtaining further reports of the aurora of 18/19 August 1950. Negative reports were received from: A. Ferreira, Lisbon (through Miss Botley), for Portugal,

VOL.
14
1959

Table 1.

Lat. N.	Long. E.	Form	Intensity	Elevation	Azimuth	Clouds	Colour	UT
								18 Aug. 1940
64° 35′	40° 31′	RB	3	45°	N	0/0	usual (sulphur)	2000–2130
65 11	49 35	HB	2	30	W–N–E	5/0	usual	2005–2105
62 14	45 01	HA	1	30	W–NW	0/0	greenish	2055–2230
65 53	34 45	C	1	90	Z	0/0	usual	2000–2140
62 52	42 43	RA	1–2	30–35	N	0/0	whitish	2132–2200
65 12	36 49	PS	1	70	N	0/0	whitish	2040–2120
65 28	39 43	C	2	15–90–15	ALL SKY	0/0	sulphur	2010–2100
64 32	40 30	RB, D	2	80, 90	W–N–E, WE	0/0	white	2000–2200
64 51	40 17	HA, D	2, 1	60, 30	W–N–E, NE	0/0	greenish	2130–2230
59 00	27 44	R	2	30	NNW–N	0/0	green/white	2015–2115
64 42	43 24	RB	2	45	N–W–E	0/0	whitish	1950–2120
63 55	45 38	RA	1, 2	30	N	0/0	light yellow	2100–2118
66 29	32 55	R	2	90	Z	8/8	vivid yellow	1950–2120
62 10	49 05	P,A R	1	40	E–NE	0/0	whitish	

Madeira and the Azores; N. Gökdogan, Istanbul, also for the Kandilli Observatory; F. Jacka, Melbourne, for Southern Victoria and Tasmania, Australia, overcast at the time (but see pp. 62–64 of the Australian National Antarctic Research Expedition Interim Report No. 5, for Observations of Aurora Australis—Macquarie Island, May 1950–April 1951); J. Paton, Edinburgh, for Scotland (then overcast) and O. Schneider, Buenos Aires, for S. America.

Positive records were reported by A. I. Lebedinsky, Moscow, who sent the attached list (Table 1) prepared by Mrs. Alsufeva of Nizmir; the observations overlapped the two 3-hour periods 18–21 and 21–24 UT on 18 August, preceding that in which the Spetsai photograph was taken; the corresponding K_p figures were 4_0 and 5-. The observations apparently did not continue after (at latest) 1 a.m. local time.

After the U.S.A. observation at Nantucket, Mass., as late as 0530 UT on 19 August, the next records of aurora are those reported by I. L. Thomsen of Wellington, New Zealand, beginning at 0747 UT. The K_p index in the intervening interval 0300–0600 UT was only 4+. The New Zealand observations began in the 3-hour interval 0600–0900 for which K_p was 6—. They may be considered as the beginning of a new and greater auroral storm extending over 19/20 August, for which the data are more numerous, but they will not be further considered here.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 123 to 137. Pergamon Press Ltd. Printed in Northern Ireland

Some measurements of horizontal movements in region-*F2* using widely spaced observing stations

L. Thomas
Department of Physics, University College of Swansea

(*Received* 10 *July* 1958)

Abstract—The results of a study of horizontal movements in region-*F2* using a system of stations spaced some 200 km apart are described. Vertical incidence recordings were used to study corresponding fluctuations at the three stations in the critical frequency and in the equivalent height of reflection at a fixed frequency.

The measured velocities ranged from 40 to 280 m/sec and the direction of travel of the perturbations was mainly towards the South-east. A limited number of simultaneous observations at a station some 650 km north of the main group showed that these large-scale irregularities travel such distances with little change in direction but there is clear evidence for larger velocities at more northerly latitudes. A comparison of the velocity magnitudes deduced from the two parameters (critical frequency and equivalent height) indicates a positive height-gradient of velocity of the order of 1 m/sec per km within region-*F2*. The results also suggest that the height gradient is larger in winter than in summer and larger near the lower edge of the layer than at the peak. Correlation, over both long and short periods, was observed between velocity magnitude and the degree of magnetic activity.

VOL.
14
1959

1. Introduction

Studies of horizontal movements in region-*F* of the ionosphere have been made by Munro (1950), Chapman (1953), Bramley (1953), Briggs and Spencer (1954) and others. Various methods have been employed but they all depend directly or indirectly on identifying and tracking some irregularity in the reflecting layer. In the experiments considered in this paper movements in region-*F2* have been studied by comparing ionograms taken at three widely spaced observatories. Most of the measurements deal with perturbations in the *F2*-layer critical frequency and consequently refer to the level of the peak of the layer but some supplementary data on drifts at a lower level were also obtained from measurements of equivalent height changes at a fixed frequency. In general, three observing stations spaced from 180 to 300 km apart were used, but a few results were also obtained for a four station network, the fourth station being some 650 km north of the main group.

2. Experimental Arrangements

In the present work ionograms obtained at three widely spaced stations, Slough, Swansea and Bangor, were examined for corresponding irregularities and from measurements of the times at which such irregularities occurred at the different stations the horizontal component of movement in region-*F2* was estimated. The relative dispositions of the three stations is shown in Fig. 1.

In some experiments ionograms were also available from a fourth station, Inverness, located some 650 km north of the Swansea-Slough baseline.

The period of observation extended from June 1951 to August 1953, but the preparation and publication of this paper have been delayed owing to the author's absence on National Service. In general, ionograms were taken for 3 days/month

at approximately 10 min intervals, but in certain months additional records were obtained continuously for a period of 1 hr before and after noon. Initially, observations were made by both day and night but due to the few irregularities observed during the night it was soon decided to limit the observations to the daylight period, approximately between 0900 and 1700 GMT. The greater frequency of irregularities in the region-*F2* during the daylight period is illustrated by the fact that in the first 15 days on which observations were made forty-four perturbations in region-*F2* suitable for the present studies were observed between the hours 0800 and 1800 GMT, whereas only twelve irregularities were observed in the remaining hours of the day.

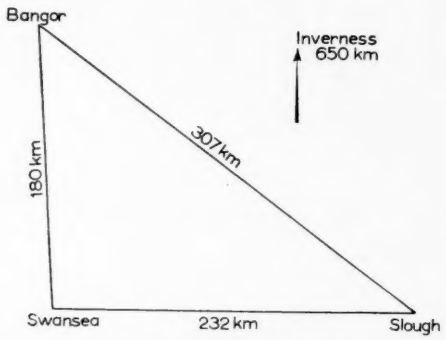


Fig. 1. Observing stations.

Two *F2*-parameters have been used to examine the horizontal movements:

(a) the ordinary ray critical frequency f_0F2,

(b) the equivalent height of reflection for the ordinary ray at 5 Mc/s;

It is clear that perturbations in the former of these will refer to the level of the peak of the layer whereas those in the latter will refer to variations in the height of a contour of constant ionization density. This contour will certainly be at a lower level than that of the peak and it was anticipated that the two sets of observations would furnish some information on drift velocities at two levels within region-*F2*. The choice of the fixed frequency for studying the equivalent height perturbations was governed by the seasonal and diurnal changes in f_0F1 and f_0F2 over the whole period of the observations.

Typical perturbations in these two parameters used in the present study are shown in Figs. 2 and 3. In Fig. 2 a perturbation in f_0F2 moving southward and observed successively at Bangor, Slough and Swansea is shown. It will be seen that variations of from 0·2 Mc/s to 0·5 Mc/s (representing changes of from 8 to 20 per cent in the peak electron density of the region) are recorded with time delays of up to 14 min at the three stations.

A typical perturbation in equivalent height at 5·0 Mc/s moving south-westwards and observed successively at the three stations is shown in Fig. 3. In general, the amplitude of the variations amounted to between 25 km and 50 km in equivalent height, but sometimes corresponding variations of smaller magnitude were also recorded at the three stations. The variation in true height of reflection would, of course, be considerably smaller than this. It is probable that some of the variations

in equivalent height arose indirectly as a result of critical frequency variations, but there were many occasions on which the variations in one parameter occurred without any measurable change in the other.

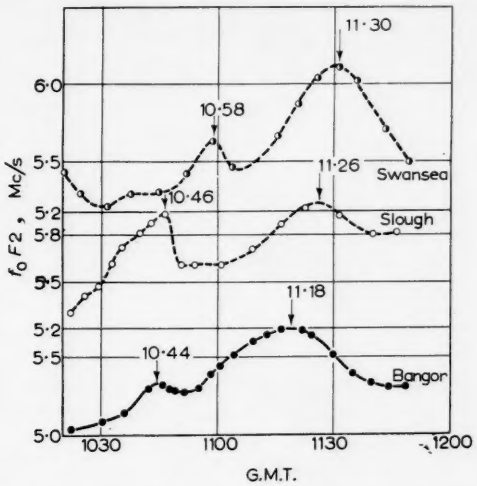


Fig. 2. Typical perturbations in f_0F2 observed at three stations.

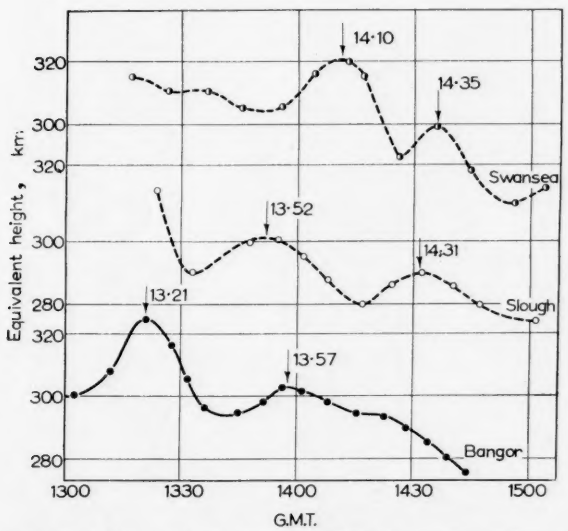


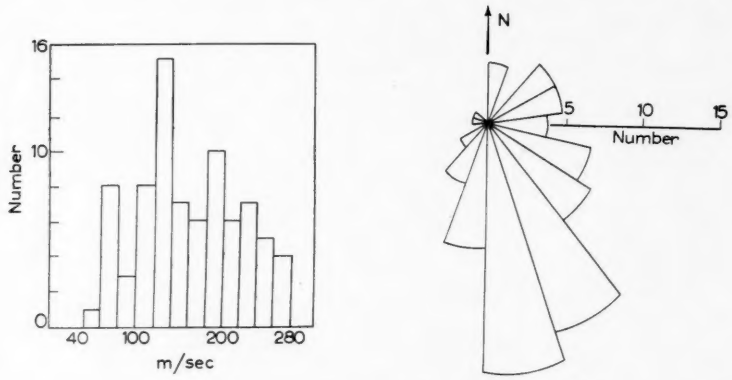
Fig. 3. Typical perturbations in h' (5·0 Mc/s) observed at three stations.

2.1. *Accuracy of the measurements*

For the purpose of estimating the time shifts between the occurrence of corresponding irregularities at the three stations all available measurements of the two parameters f_0F2 and h' at 5·0 Mc/s were plotted against time and smooth curves drawn through the points. In many cases measurements were only available

VOL. 14 1959

at one or other of the stations at intervals of from 10 to 15 min, and this gave rise to some uncertainty in the determination of a maximum or minimum in the plotted curves. However, for a number of days, measurements were made continuously at all three stations over a period of 2 hr (1 hr in the morning and 1 hr in the afternoon) and we have used these continuous measurements to form some estimate of the inaccuracy likely to arise in the velocity vectors when calculated from measurements at 10–15 min intervals. We find that in general this error due to uncertainty in fixing the times of maxima or minima may be expected to be



(a) (b)

Fig. 4. Distribution of drift velocities based on f_0F2 measurements: (a) magnitude, (b) direction.

about ± 30 m/sec in the magnitude of the vector and to about $\pm 10°$ in the calculated direction.

It is to be added that for various reasons it was found necessary to reject a number of cases in which clear correlation existed. These included cases in which the irregularities occurred in very rapid succession at all the observing stations indicating very large apparent velocities of horizontal movements (of the order of 500 m/sec and larger). Other cases excluded were those in which an irregularity appeared to have undergone such a marked change of form during its travel between the widely spaced observing points as to make correlation uncertain.

3. Experimental Results

3.1. *Drift velocity based on measurements of f_0F2*

3.1.1. *Magnitudes.* In the course of these experiments eighty reliable values of drift velocity deduced from irregularities in f_0F2 were obtained and Figs. 4(a) and 4(b) show the distribution of velocities (magnitude and direction) obtained. It will be seen that the magnitudes are distributed over the range from 40 m/sec to 280 m/sec with most frequent values in the range from 120 to 140 m/sec. The median value is 155 m/sec.

The direction of movement shows a predominance of drift towards the south with about 60 per cent of the values in the SE quadrant. Very few drifts were observed in the SW–NE sector.

3.1.2. *Seasonal changes in velocity.* Although the observation programme extended over the 2 year period June 1951 to August 1953 it was only from February 1953 that observations were possible at monthly intervals. Hence, in examining the present data for evidence of a seasonal variation, the results for months of different years have been interposed with each other to give data for a complete year. The results are given in Figs. 5(a) and 5(b) and include all measurements in the period 0900–1700 GMT. Fig. 5(a) shows that the magnitude of the drift vector varied considerably from month to month and in these data there appears

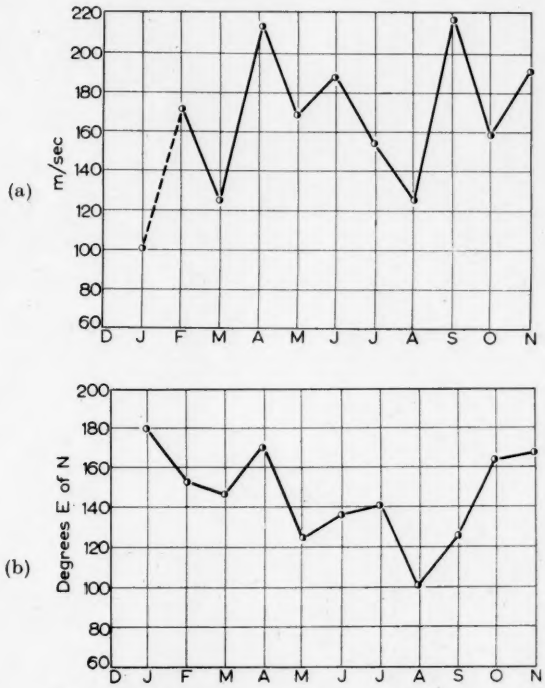


Fig. 5. Seasonal variation in velocities (median values): (a) magnitude; (b) direction.

to be little significant evidence for a seasonal variation. However, if noon data alone are considered there is some evidence indicating that the average velocity magnitude is larger in winter than in summer (see below). Fig. 5(b) shows that the direction of movement varied from about 150°–170°E of N in winter to about 120°–140°E of N in summer.

3.1.3. *Diurnal changes in velocity. Magnitude.* The diurnal variation in the magnitude of the drift vector is shown in Fig. 6(a). It will be seen that in both winter and equinox the drift velocities are at a maximum around noon but in summer the data indicate a minimum near noon.

Direction. (Fig. 6(b).) In winter the direction of drift shows a gradual anticlockwise rotation from about 200° E of N near 1000 GMT to about 60° E of N in the late afternoon. There is some evidence for a reversal of this tendency in the early evening.

At the equinox period there appears to be no significant variation in direction between 0800 and 1800 GMT the drift always being in the 160°–180° sector. In the late afternoon there is again some evidence for a clockwise rotation of the vector towards the south-west.

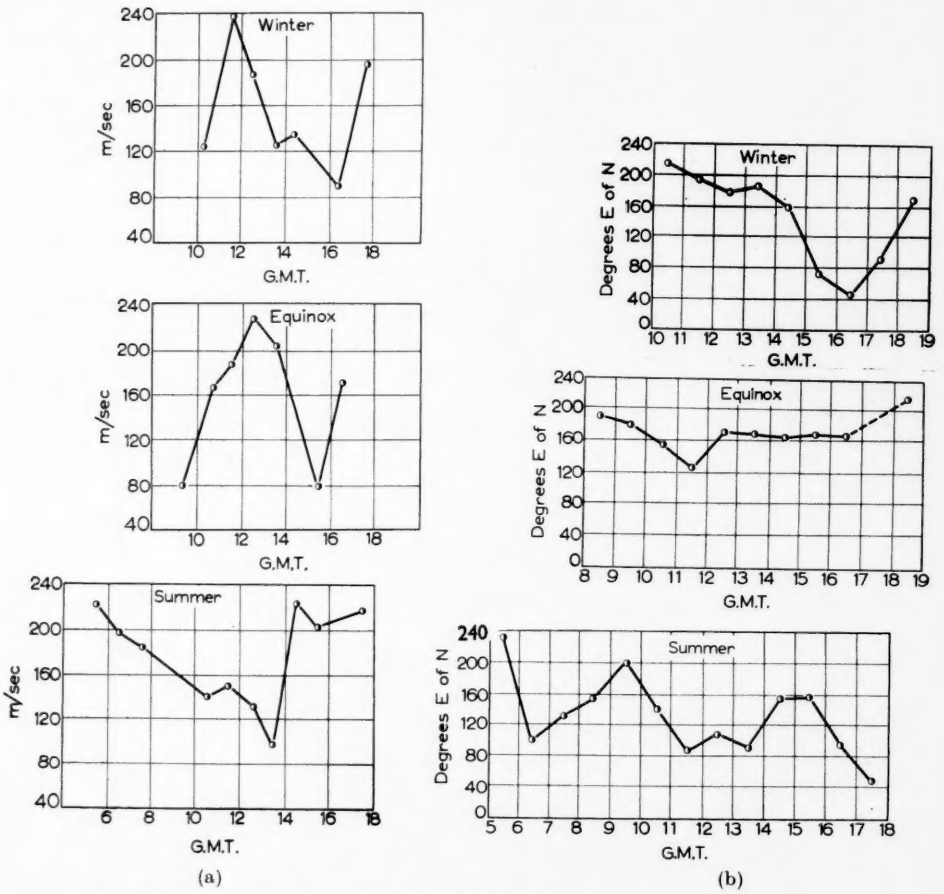


Fig. 6. Diurnal variation in velocities in winter, equinox and summer: (a) magnitude; (b) direction.

In summer there is little variation in direction over the day, the mean direction being about 140° E of N.

3.1.4. *Influence of magnetic activity.* In Fig. 7 the magnitude of the drift velocity over the period 1000–1400 GMT is plotted against the magnetic disturbance index K as measured at Abinger and Eskdalemuir. It will be seen that for values of K less than 2·5 there is no evidence for a dependence of the velocity on the magnitude of K but for larger disturbances there is a definite increase of velocity with increased activity. A correlation with magnetic activity is also shown in Fig. 8 in which the monthly median velocities and magnetic activity indices K are plotted.

The K-magnetic indices are normally only published for 3 hr intervals and it is of interest to examine the existence or otherwise of detailed correlation between drift velocity and magnetic disturbance within a 3 hr interval. Unfortunately, the intervals over which drift measurements were available were seldom periods

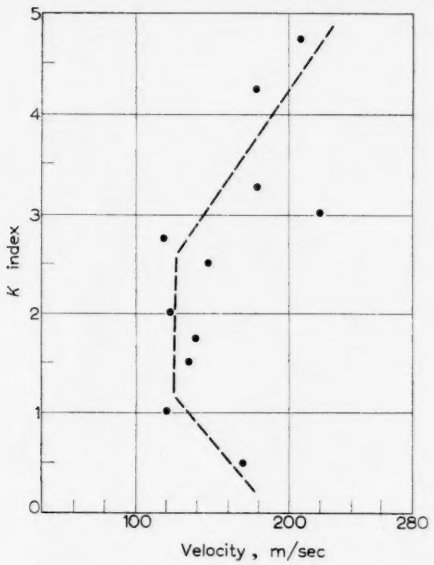


Fig. 7. Velocity magnitude and magnetic disturbance index K (mean value for Abinger and Eskdalemuir).

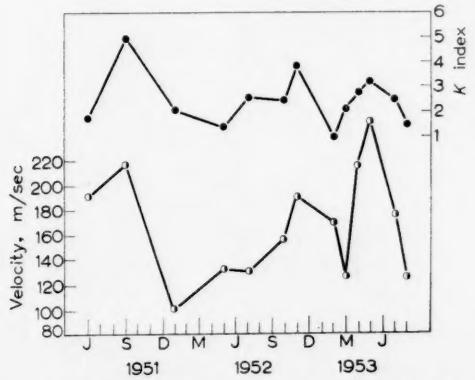


Fig. 8. Monthly median drift velocity and magnetic disturbance index K (mean of Abinger and Eskdalemuir).

of great magnetic disturbance but we were able to select a few periods in which a measurable change in magnetic activity and in velocity occurred within a period of hours. Two examples will be quoted here. In each we have measured on the original magnetograms the range in the horizontal component of the earth's field over half-hourly or hourly intervals. Fig. 9(a) shows the variation in velocity and

horizontal range (R_H) over a magnetically quiet period of about 8 hr. It will be seen that although the 3 hr magnetic index remains constant at $K = 1$ there was a measurable perturbation in H amounting to about 15 γ which coincided in time with a large variation in the drift velocity. Fig. 9(b) shows another example in which K varied from 1·5 to 3 with a maximum within a 9 hr interval. In this K showed a slight maximum in the 3 hr interval 0900–1200 GMT whereas the more

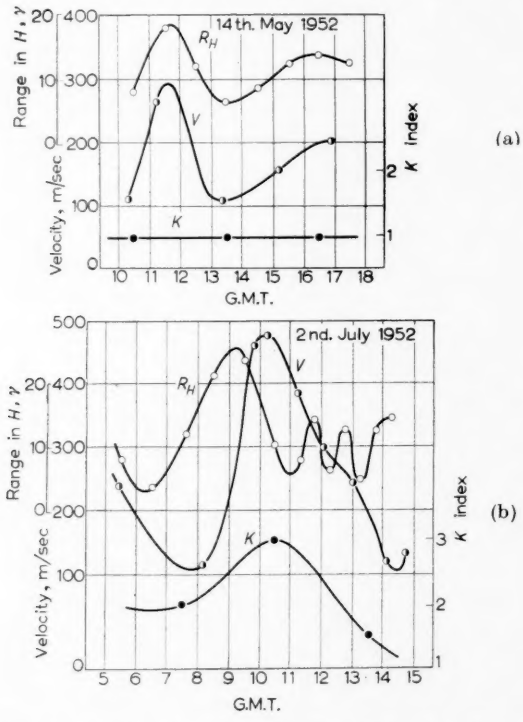


Fig. 9. Examples of perturbations in magnetic field and in drift velocity over periods of a few hours.

detailed variation in activity, as shown by R_H, showed that the actual maximum in activity occurred at the beginning of the interval.

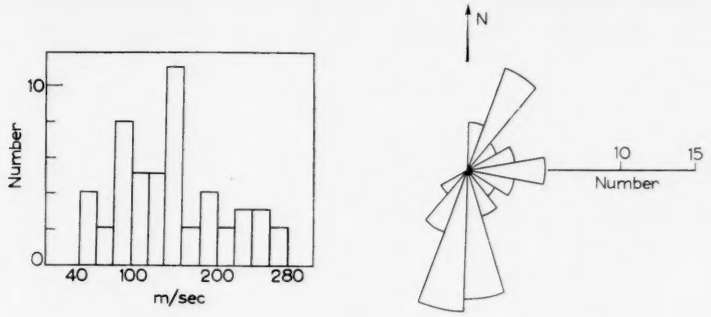
3.2. *Drift velocity based on measurements of h' at 5·0 Mc/s*

Fig. 10(a) shows the distribution of velocity magnitudes for fifty-one measurements of equivalent height perturbation on the fixed frequency 5·0 Mc/s. It will be seen that the overall range of velocities observed is about the same as that found for the critical-frequency observations but the distribution indicates a rather smaller proportion of large values. The median velocity is 142 m/sec (compared with 155 m/sec for the set of critical-frequency measurements).

The direction of drift (Fig. 10(b)) shows a complete absence of values in the NW quadrant. About half the values are in a southerly direction with a smaller but significant proportion towards the NE.

Although the number of fixed-frequency drift measurements was considerably smaller than that obtained from critical-frequency perturbations it is of interest to note the order of magnitude of the average velocities found in the daylight period in the three seasons. These were: summer 135 m/sec, equinox 107 m/sec and winter 165 m/sec. (The winter figure is based on only 2 days measurements and is less reliable than those for summer and equinox.)

There were insufficient numbers of observations to yield any significant information on a possible diurnal variation either in magnitude or direction of the drift velocities but we were able to confirm that in these measurements, as with



(a) (b)

Fig. 10. Distribution of drift velocities based on measurements of $h'(5{\cdot}0$ Mc/s): (a) magnitude; (b) direction.

those on critical-frequency changes, there was a significant positive correlation between magnetic activity and drift speed.

3.3. *Observations over an extended base line*

At certain periods during these experiments frequent ionospheric soundings were also made at Inverness, a station some 650 km north of the Swansea–Slough baseline. A comparison of the data from the four stations showed that only on rare occasions could definite correlation in equivalent height perturbations be detected at all four stations but that a number of examples of correlation between comparatively large variations in f_0F2 could be detected. Over a period of 10 days some nineteen reliable results were obtained for this four station system. For these nineteen results it was found that the median velocity deduced from the data for the three southerly stations (Bangor, Slough, Swansea) was 179 m/sec in a direction 146° E of N. If data from Inverness, Slough and Swansea were used the mean velocity was found to be 235 m/sec in a direction 151° E of N. Thus it would appear that these large scale disturbances which appear at all four stations move southward in a fairly constant direction but that the velocity of movement is larger at northern latitudes and decreases considerably as they come towards the south.

4. Discussion of Results

Long series of measurements on *F2*-layer drifts with which the present results can be compared have been made by Phillips and Spencer in England and by Munro in Australia.

VOL. 14 1959

PHILLIPS and SPENCER used the closely-spaced-receiver fading method on frequencies between 2 and 4 Mc/s and a summary of their results for the period 1949–1953 has been published in the survey report by BRIGGS and SPENCER (1954). As BRIGGS and SPENCER point out, results obtained by the fading method on movements in region-F may sometimes be open to doubt especially during the summer day-time when the sounding frequency (4 Mc/s) is not much greater than the E-region critical frequency. However, for this present discussion we shall consider the measurements as given in the published report, the possibility of uncertainty in the values for the summer day-time period being borne in mind.

The southern hemisphere measurements of MUNRO involved the observation of irregularities in equivalent height using three fixed-frequency transmissions on a frequency of 5·8 Mc/s. The separation of the transmitters was of the order of tens of kilometres whereas in the present experiments the observing points were some 200–300 km apart.

4.1. *Magnitude of velocity*

The median drift speed for eighty observations in the daylight period in the present experiments is 155 m/sec. These measurements will refer to the level of the peak of the $F2$-layer and in this connexion it is relevant to note that the average value of f_0F2 for the period of the measurements was about 6·5 Mc/s. The median velocity from the 5·0 Mc/s equivalent height measurements is a little smaller, viz. 142 m/sec. The corresponding median value deduced from the Cambridge results on 4 Mc/s is 37 m/sec and that given by MUNRO's observations on 5·8 Mc/s is 128 m/sec. The median values found in the present experiments are thus somewhat larger than that given by MUNRO and are several times that given by the fading method. In making this comparison between the three sets of results the following points have to be noted:

(a) Different methods of measurement are involved and the spacing between the observing stations varies from a few hundred metres to some hundred kilometres.

(b) Two sets of results refer to latitude 51°N whereas the third refers to latitude 34°S.

(c) The measurements refer to different levels in the F-layer since different frequencies have been used.

Although the points of difference (a) and (b) are not to be ignored, there are strong grounds for believing that the really significant feature is (c) and that the overall results probably indicate a marked positive gradient of velocity in the F-layer of the ionosphere.

Support for this view is given when data for the separate seasons are considered. Table 1 shows the mean day-time velocities in the three seasons, summer, winter, equinox for each set of measurements.

It will be seen that in the winter and equinox seasons the larger velocities are found on the high frequencies. In summer the average value of f_0F2 at the time of the Swansea experiments was quite near the frequency used by MUNRO and in this season the drift magnitudes are also equal. It will also be noted that for all the sets of results winter velocities are larger than those in summer, but that only

in the case of the 4 Mc/s results do the equinox values lie between those for summer and winter.

In the absence of precise information on ionospheric conditions at the time of the various measurements only an approximate estimate can be made of the height gradient of velocity which these measurements indicate. Rough estimates of the

Table 1. Average velocity (m/sec) over period 1000–1400 GMT

Deduced from	Summer	Equinox	Winter
Cambridge 4 Mc/s (fading)	25	36	80
Swansea 5 Mc/s (equivalent height)	135	107	*
Sydney 5·8 Mc/s (equivalent height)†	130	110	133
Swansea (critical frequency)	130 (5·4 Mc/s)	180 (6·5 Mc/s)	170 (7·2 Mc/s)

* Insufficient data available.
† Results for Sydney refer to time interval 0900–1600 Eastern Australian Time.

VOL. 14 1959

true levels of reflection of the various frequencies involved have been made and Table 2 shows the estimated gradients based on the Cambridge, Sydney and Swansea results.

Great reliance is not to be placed on the magnitude of the winter value of 4·4

Table 2. Approximate estimates of height gradient of velocity (m/sec per km)

Deduced from	Summer	Winter
Cambridge (4 Mc/s) and Sydney (5·8 Mc/s)	2	4·4
Cambridge (4 Mc/s) and Swansea (critical frequency)	1·1	1·6

m/sec per km since this is based on a comparatively small height difference but these figures suggest that over the height ranges considered (180–250 km) there is a larger gradient of velocity in region-*F2* in winter than in summer and also a smaller gradient towards the peak of the layer than near the lower edge. An additional estimate of the order of magnitude of the height gradient of velocity is provided by comparing velocities based on critical frequencies measurements with those deduced for the same days from equivalent height measurements at 5·0 Mc/s. Such a comparison indicates an average gradient of about 1 m/sec per km for a sample of fifty measurements spaced over both summer and winter months.

4.2. *Direction of drift*

The published work of MUNRO (1950, 1953) reveals two or three significant features concerning the direction of drift in region-*F2* as observed at Sydney.

(a) Throughout the year the mean direction of drift in the day-time is towards the east—in every month the median direction lies in the sector between 20° and 150° E of N.

(b) There is a marked seasonal change in direction with a swing (occurring abruptly at the equinoxes) from 20° E of N in local winter to about 110° E of N in local summer.

(c) In local winter there is a steady rotation in the drift vector from about 50° E of N at 0900 Eastern Australian Time to about 15° E of N near 1500 EAT. MUNRO also states that there is some evidence in his results for a significant change of direction from towards the east in the day to towards the west at night.

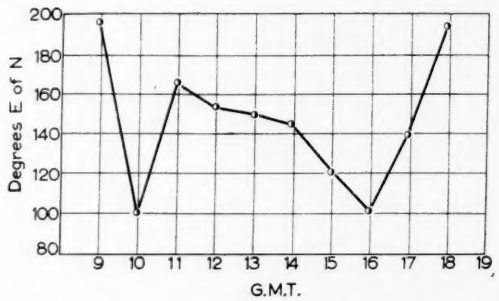


Fig. 11. Diurnal variation in direction of drift velocity at Cambridge (4 Mc/s—fading method).

Examination of Fig. 5(b) shows that both the general direction of drift and seasonal change in direction found by MUNRO are also evident in the present results. Thus the median direction always lies between 100° and 180° E of N and whereas the mean direction for the months of November, January and February is 168° E of N, the mean for the months June, July and August is 126° E of N, representing a swing of 42° towards the *north* from local winter to summer. MUNRO's result gave a swing of 90° towards the *south* from local winter to summer.

The diurnal variation in direction in winter is also found in the present results. Fig. 6(b) shows that the drift vector rotates from 200° E of N near 1000 GMT to about 50° E of N between 1600 and 1700 GMT, a rotation of 150° towards the north as the day proceeds. After 1700 GMT there is some evidence for a rapid swing back towards the south and this might well fit in with the suggestion that at night the direction of drift is towards the west.

Fig. 11 shows the diurnal variation in direction in winter observed on 4 Mc/s by the fading method and comparing this with Fig. 6(b) shows that there is quite remarkable agreement between the two sets of measurements. Both curves show the steady swing of the drift vector towards the north between 1100 and 1600 GMT followed by a rapid swing back towards the south between 1600 and 1800 GMT. Between 1800 and 0900 GMT the 4 Mc/s results also show that the drift is towards the west as suggested by MUNRO's results.

VOL
14
195

The present results do not show any well defined diurnal variation in the direction of drift in summer between 0730 and 1600 GMT but there is some suggestion of a change in direction from about 230° E of N near 0500 GMT to about 40° E of N near 1800 GMT. The only published values with which these can be compared are those from the 4 Mc/s fading method. These show a marked swing from 280° to 30° near sunrise with a rotation back to the original direction again near sunset.

4.3. *Drift velocity and magnetic activity*

The present work on the movement of large-scale irregularities in region-*F2* shows that there is some positive correlation between the magnitude of the drift velocity and the degree of magnetic activity. For all the available data the correlation coefficient between mean velocity magnitude and K_p (the planetary index of magnetic activity) is $0{\cdot}65 \pm 0{\cdot}2$. Although Fig. 7 shows that the drift

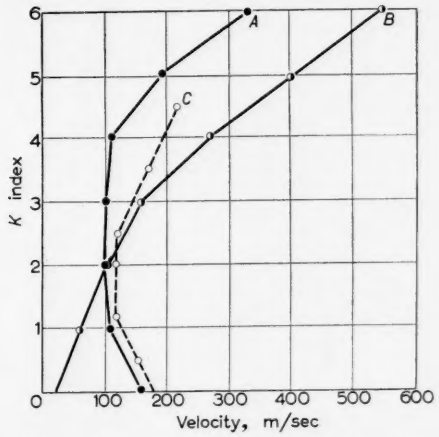


Fig. 12. Velocity magnitude and magnetic disturbance index K. Curve A: Cambridge results on 4 Mc/s. Curve B: Manchester radio star studies. Curve C: Swansea f_0F2 perturbations.

velocity does not appear sensitive to changes in K when this index is less than about 2·5, we have found that in some cases quite small perturbations in the magnetogram trace coincided with enhanced velocities. Further evidence for the dependence of velocity on magnetic activity is provided by the results from the four station network in which the velocities based on ionograms from the northernmost station (Inverness) were distinctly larger than those based on the more southerly stations. This might be interpreted as showing that higher velocities occur at latitudes nearer the zone of auroral and magnetic activity. A similar result has been found in studies of radio star scintillations (MAXWELL, 1954). Radio star scintillation studies also show the general dependence of velocity magnitude on magnetic activity but the velocities of the irregularities detected in these methods (located near the 400 km level) appear to be more sensitive to disturbance than is found from ordinary radio sounding experiments.

Fig. 12 shows three experimental curves relating drift velocity to magnetic

disturbance index K at this latitude. Curve A refers to F-layer drift measurements at Cambridge by the fading method using a frequency of 4 Mc/s. Curve B is that obtained in radio star studies (MAXWELL, 1954) and curve C is that obtained from the present studies on the movement of irregularities at the peak of the *F2*-layer. It will be seen that for the high-level irregularities responsible for radio star scintillations there is a measurable increase in velocity even for the smallest changes in activity. In the case of irregularities near the peak of the *F2*-layer (curve C) the response begins for values of K above about 2·5 but the increase is not as marked as for the scintillation irregularities. For the 4 Mc/s drift results (which refer to a considerably lower level in the F-layer), there appears to be no response for values of K less than 4·0. At much lower levels in the ionosphere, measurements on E-layer drift velocities show that the velocity is only increased in periods of severe magnetic activity. The experimental facts for temperate latitudes thus appear to indicate a greater sensitivity of drift velocity to magnetic

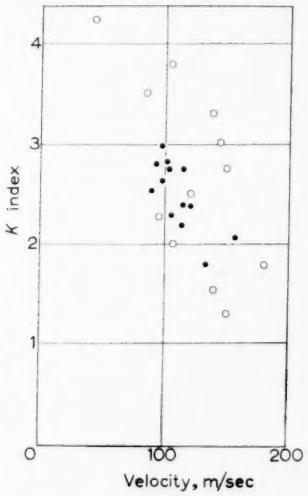


Fig. 13. Velocity magnitude and magnetic disturbance deduced from MUNRO's results showing some tendency for negative correlation. ● Monthly median velocities 1948–1949. ○ Daily values for November 1948.

activity with increasing altitude. Measurements by CHAPMAN (1953) using the fading method at Ottawa, a station nearer the auroral zone than those considered above, shows that in the E-layer there is again little response to activity if K is less than 4·5 but that F-layer velocities increase regularly from the lowest values of K.

MARTYN (1947, 1948) has shown that horizontal movements in *F2* arise as a result of electro-dynamic forces which have their origin in the lower "dynamo" region. It is thus not unexpected that there should be some correlation between drift velocity and magnetic activity. However, the different degree of sensitivity which is apparent at different levels suggests that in the E- and F-layers there are probably additional factors influencing drift velocity but that at the highest levels the governing factor may be more completely electrodynamic in character.

In his discussion of F-layer drift measurements at latitude 34°S, MUNRO (1950) makes no reference to any correlation with magnetic activity but PRICE (1954) states that for measurements at 30°S no correlation between drift velocity and

magnetic activity could be detected. We have re-examined some of the experimental data published by MUNRO and this would appear to show that there is a small but definite *negative* correlation between velocity magnitude and magnetic activity. Fig. 13 shows a plot of velocity magnitude measured by MUNRO near Sydney and magnetic disturbance index-K measured at the Toolangi Observatory (about 600 km from Sydney). Two sets of points are plotted. The first refer to monthly median values measured by MUNRO between April 1948 and April 1949. The second set have been deduced from a set of daily values given by MUNRO for the month of November 1948. It will be seen that both sets of measurements suggest a weak negative correlation between velocity magnitude and magnetic activity, at least for the limited range of K from 1·5 to 4·0. Unfortunately, no data were available for conditions of great magnetic disturbance but it seems possible that the variation shown in Fig. 13 may perhaps correspond to the lower parts of curves A and C (Fig. 12), both of which show some degree of negative correlation for low magnetic activity.

In the course of this work a few examples were noted of detailed correlation over a period of a few hours between drift velocity and magnetic activity. In all cases the correlation observed was between the velocity and R_H the range in the horizontal component of the magnetic field and maxima in the one were accompanied by corresponding maxima in the other.

VOL. 14 1959

Acknowledgements—The author wishes to acknowledge valuable discussions with Dr. W. J. G. BEYNON and his assistance in the preparation of this paper.

During the period of this work the author was in receipt of a D.S.I.R. studentship.

This work was carried out as part of the programme of the Radio Research Board and this paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

REFERENCES

BRAMLEY E. N.	1953	*Proc. Roy. Soc.* A **220,** 39.
BRIGGS B. H. and SPENCER M.	1954	*Rep. Progr. Phys.* **17,** 245.
CHAPMAN J. H.	1953	*Canad. J. Phys.* **31,** 120.
MARTYN D. F.	1947	*Proc. Roy. Soc.* A **189,** 241.
	1948	*Proc. Roy. Soc.* A **194,** 445.
MAXWELL A.	1954	*Physics of the Ionosphere* p. 166. Physical Society, London.
MUNRO G.	1950	*Proc. Roy. Soc.* A **202,** 208.
	1953	*Nature, Lond.* **171,** 693.
PRICE R. E.	1954	*Physics of the Ionosphere* p. 181. Physical Society, London.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 138 to 166. Pergamon Press Ltd. Printed in Northern Ireland

Geomagnetic distortion of region-E

W. J. G. Beynon and G. M. Brown
University Colleges of Aberystwyth and Swansea
Official Communication from DSIR Radio Research Station, Slough

(*Received* 7 *September* 1958)

Abstract—A detailed consideration of the behaviour of the normal E-region shows the existence of minor perturbations from the classical Chapman theory. It is shown that some of these differences can be ascribed to the influence of vertical drift of ionization resulting from the interaction of the earth's magnetic field and the S_q-current system flowing in or near the E-layer. The resulting distortion of the electron density profile of the E-layer is considered, with particular reference to its effect on the values of the parameter f_0E at different latitudes. The diurnal and seasonal changes in the magnitude of the east–west component current are clearly reflected in the variations of f_0E, and an index n, measuring the sensitivity of the E-layer to seasonal changes, shows marked singularities near the latitudes of the S_q-foci.

An analysis of conditions on magnetically disturbed days is also given, and it is concluded that at temperate latitude stations there is a depression in fE at times of magnetic disturbance comparable with the magnitude of that associated with the vertical drift effect of the S_q-current system at noon in the summer. Reference is also made to conditions near and within the auroral zone. Consideration of the effect of sporadic E-ionization shows that this does not appear seriously to influence the tabulated values of fE.

1. Introduction

There is now abundant experimental evidence that the ionization of the normal region-E of the ionosphere is closely governed by solar ultra-violet radiation. The diurnal and seasonal variations in the critical frequency of this layer are remarkably regular, and the long-term changes are also closely associated with corresponding changes in the intensity of the incoming radiation. In recent years, however, careful measurements of E-layer characteristics have revealed the existence of certain anomalies between experimental data and results expected from the usual Chapman theory of a solar-produced ionized layer.

Hitherto, studies of the world pattern of ionospheric phenomena have largely been confined to region-$F2$, partly because the various anomalies of that region are clearly discernible in the published data and partly because of the importance of region-$F2$ in practical radio communication problems. However, the work of Martyn (1952) and others has indicated that in some important respects phenomena in region-$F2$ may be profoundly affected by events in region-E. In particular, it is now believed that region-E is the seat of the electric currents which flow in the upper atmosphere and are responsible for the diurnal quiet-day S_q-variations of the magnetic elements, and it has become important to analyse critically the published data for this region. In this paper we discuss the behaviour of the E-layer under magnetically quiet and disturbed conditions, and conclude that this shows clear evidence for the perturbing influences of the magnetic current systems. Although there are certain anomalies of the E-layer for which the explanation is not at present clear, there is little doubt that this geomagnetic distortion of the layer is responsible for several of the reported discrepancies between observation and theory. The analysis presented here is mainly confined to an examination of critical frequencies, and a summary of the relevant theoretical relationships is first given.

VOL
14
195

2. Some Theoretical Considerations

2.1. *General*

For a simple solar-controlled region, the variation of the electron density N with time t is governed by an equation of the form

$$\frac{\partial N}{\partial t} = q - \alpha N^p - \text{div}\,(N\boldsymbol{v}) \tag{1}$$

where q is the rate of electron production per unit volume at the level concerned, α is the effective recombination coefficient and $\boldsymbol{v}$ is the vector drift velocity. The third term on the right-hand side of this equation, expressing the influence of electron transport, is, in general, smaller than the other two terms. The value of the index p depends on the processes of electron dissipation in the region.

The rate of electron production has a maximum value q_m given by

$$q_m = q_0 \cos\chi \tag{2}$$

and occurs at a height

$$(h)_{q_m} = h_0 + H\,(\ln \sec\chi) \tag{3}$$

where χ is the solar zenith angle, q_0 is the rate of electron production per unit volume for overhead sun conditions, occurring at a height h_0, and H is the scale height of the atmosphere at the level concerned. When χ exceeds about 75°, sec χ must be replaced by a function $Ch(\chi)$, tables of which have been given by Chapman (1931, 1953). Owing to changing solar activity the intensity of the incoming solar radiation varies, and we note that q_0 is related to S_∞ (the intensity of the solar radiation outside the atmosphere) by the relation

$$q_0 = \frac{\beta S_\infty}{He} \tag{4}$$

where β is the number of electrons produced by absorption of unit energy and "e" is the exponential. If β and H are assumed constant, then q_0 is proportional to S_∞. The varying distance of the earth from the sun causes S_∞ to be approximately 6 per cent greater in January than in July.

2.2. *Noon conditions*

Near noon, the ionization density is in equilibrium with the ionizing intensity, i.e. $dN/dt \fallingdotseq 0$, and if, in addition, the movement term div $(N\boldsymbol{v})$ is small, it may be assumed that in the E-layer the levels of maximum production of ionization and of maximum density are practically the same. In this case equations (1) and (2) give the approximate relation

$$q_0 \cos\chi = \alpha {N_m}^p \tag{5}$$

where N_m stands for the maximum electron density. The latter is related to the ordinary ray critical frequency f by the expression

$$N_m = kf^2 \tag{6}$$

where $k = 1{\cdot}24 \times 10^4$ when f is in Mc/s. Equation (5) is thus

$$q_0 \cos\chi = \alpha k^p f^{2p} \tag{7}$$

VOL. 14 1959

In general, we shall write equation (7) in the form

$$f^n = A \cos \chi \tag{8}$$

so that $n = 2p$ and $A = (q_0/\alpha)k^{-n/2}$. Equation (8) can also be written

$$f = f_s \cos^{1/n} \chi \tag{9}$$

where f_s is the sub-solar critical frequency, and is given by

$$f_s = A^{1/n} = \left(\frac{q_0}{\alpha}\right)^{1/n} k^{-(1/2)} \tag{10}$$

It is to be noted that f_s depends on q_0, α, and on the value of the index n.

2.3. *The effect of gradients of H and* α

If the scale height H is not constant, but has a height gradient λ (for example if the atmosphere is not isothermal, or the mean molecular mass is not constant), then it can be shown that equation (8) takes the modified form

$$f^n = A_1 (\cos \chi)^{1+\lambda} \tag{11}$$

and equation (9) becomes

$$f = f_s (\cos \chi)^{\frac{1+\lambda}{n}} \tag{12}$$

A similar modification to the simple expressions (8) and (9) results from consideration of the effect of a possible height gradient of recombination coefficient α.

2.4. *The effect of vertical drift*

Martyn (1947a,b, 1948) has shown that important perturbations of an ionized layer may arise from the *vertical component* of the movement term div ($N\boldsymbol{v}$). The effect of vertical drift in distorting the electron density profile of Chapman-like layers has also been discussed by Kirkpatrick (1948) and by Appleton and Lyon (1954). Martyn has concluded that in region-E the distortion due to vertical drift is unlikely to be large, since in this region the lifetime of the ions is comparatively small and the loss of ionization due to transport is offset by the normal replenishment by solar radiation. He has shown that, under typical conditions, vertical drift in the E-layer may change the peak density by a few per cent.

Appleton and Lyon have shown that the fractional change in the peak electron density, $\Delta N_m/N_m$, of an ionized stratum of E-layer type due to vertical drift effects is given approximately by the expression

$$\frac{\Delta N_m}{N_m} = -\frac{1}{2\alpha N_m}\frac{\partial v}{\partial h} - \frac{1}{4H^2}\left(\frac{v}{2\alpha N_m}\right)^2 \tag{13}$$

and that the height of the peak electron density changes by the amount

$$(\Delta h)_{N_m} = \frac{v}{2\alpha N_m} \tag{14}$$

where H is the scale height in the layer and v is now the *vertical* velocity of electron drift, reckoned positive upwards. It will be seen from the second term on the

right-hand side of equation (13) that the change of the maximum ionization density is independent of the *direction* of drift (upwards or downwards). If there is negligible height-gradient of the drift velocity, or if there is a positive gradient $\partial v/\partial h$, distortion due to vertical drift may always be expected to *reduce* the peak density. The direction of drift will determine whether the height of the electron peak rises or falls (equation (14)), but in either case the ionization will be spread out vertically resulting in a diminished peak density. The effect of a negative velocity gradient will be to compress the ionization vertically tending to give an enhanced peak density, but the drift velocity itself will tend to remove this peak to another height, which will no longer be the height of maximum electron production. Whether the net result is an increase or decrease in N_m clearly depends on the relative magnitudes of the two terms in expression (13). It should be noted also that if either H or α varies significantly over the vertical extent of the "E-region" (Section (2.3)), then the peak electron density will, from this cause alone, vary with the height at which the peak occurs.

3. The Experimental Data

In attempting an analysis of the published data for region-E in relation to the theoretical considerations outlined in Section (2) above, we are at once confronted with certain difficulties and these will now be considered.

VOL. 14 1959

3.1. *The parameter fE*

In most of the analysis which follows we shall be concerned with the ordinary-wave critical frequency of the E-region, f_0E. Values of f_0E are, in general, measured only to the nearest 0·1 or 0·05 Mc/s, and the monthly median values are tabulated to the same order of accuracy. It is clear that an analysis of a considerable amount of data will be necessary in order to establish the reality of small perturbations of critical frequency. Moreover, the characteristics of the ionospheric recorders, and the types of "ionograms" produced differ considerably from station to station, and consequently when data are analysed for a variety of stations a certain inhomogeneity must arise. Furthermore, the procedures adopted for the scaling of the parameter f_0E have only recently been regularized.

In practice, it is found that only on comparatively few occasions does the "E-trace" show an unambiguous critical frequency cusp of the type familiar for the F-trace. The normal E-layer is very frequently stratified, and it sometimes becomes difficult to distinguish "abnormal" layers from "normal" layers. Particular difficulties in scaling fE arise when traces due to "sporadic-E" ionization are present, or when there is heavy absorption in the vicinity of the critical frequency. In such circumstances the normal E-trace may become prematurely cut-off or obscured, and if no attempt is made to extrapolate the trace it might be anticipated that scalings of the cusp-frequency as fE will lead to erroneously low values.

In this connection, attention may be drawn to the work of Bibl (1951), who has considered the question of the scaling of the critical frequency of the E-region. He has proposed a somewhat arbitrary distinction between so-called normal and abnormal E-layers, and has also given a clear definition of the parameter fE which

when applied to data for Freiburg was found to reduce the variability considerably. It seems clear that with special effort an improvement in the quality of fE data can be effected, but hitherto there has been no agreed standardization of scaling practice.

Before proceeding to a detailed examination of experimental fE data, it is clearly imperative to discover whether or not the presence of E_s-ionization has markedly affected the tabulated values. In view of the importance of this, we have critically examined certain fE data for E_s-contamination and the results of this study are outlined below.

3.2. *The influence of E_s ionization on values of fE*

Tables of published data show that E_s-ionization occurs so frequently at hours near noon, and especially during summer months, that it is not possible to consider an analysis confined exclusively to days on which E_s is absent. Examination of data shows that there are only a very few days when fE_s is lower than fE, and this number is certainly too small to allow the values of fE on these days to be taken as representative of the true values. However, we can proceed with our analysis if we make use of a fact first noted by Appleton *et al.* (1939) that the greater the intensity of E_s-ionization, the lower is its height of occurrence. Hence it might be expected that the most intense E_s-ionization will give rise to the largest errors in scaling fE.

In our first examination we selected a station (de Bilt, 52·1°N, 5·2°E) for which both values of fE_s and of $h'E_s$ were available to us. Fig. 1 shows the results of a superposed–epoch analysis of the $h'E_s$ data for all hours of the day over 1954–1957, covering an 11 day period centred on a "zero-day" selected according to the time of occurrence of intense E_s-ionization. In choosing the zero-day, an arbitrary criterion was used, viz., $fE_s > 5$ Mc/s. It is clear that there is a definite tendency for the more intense E_s-ionization to occur at a lower height, in agreement with the result of Appleton *et al.* (1939). A similar analysis of the published f_0E data for the same period gives the lower curve of Fig. 2. The upper curve shows the mean value of fE_s for the same days. It would appear that there is a depression of about 0·03 Mc/s on the day of very intense E_s-ionization, but on the other hand, there are also perturbations of order $\pm$0·03 Mc/s on at least three of the other days, and these certainly cannot be associated with abnormal changes in the intensity of E_s-ionization.

In our subsequent work we shall have frequent occasion to refer to fE data for the stations Slough and Tokyo over the period 1952–1954. We have therefore carried out similar superposed–epoch analyses for these stations. Fig. 3 shows the results for Slough, using noon fE values (lower curve), and taking $fE_s \geqslant 7$ Mc/s (upper curve) as a criterion of the most intense E_s-ionization (zero-day). For Tokyo (Fig. 4) it was possible to find a sufficient number of values of fE even on days when $fE_s \geqslant 9$ Mc/s. In the case of Fig. 3 it is seen that the perturbations in fE are of the same order as those in Fig. 2, and that on the day of maximum fE_s ($\sim$8·8 Mc/s) the value of fE is possibly 0·03 Mc/s below average, but here again variations of the same magnitude occur on a number of the other days considered without any obvious associated change in fE_s. For Fig. 4, a mean level of 10 Mc/s

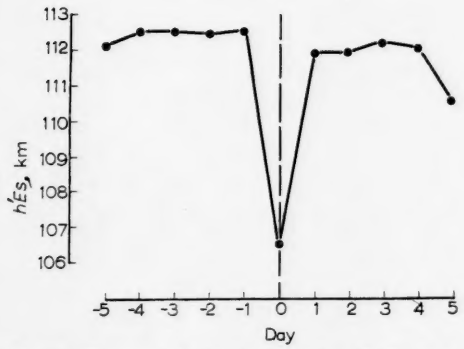


Fig. 1. Superposed–epoch analysis of $h'E_s$ data for de Bilt, all hours 1954–1957. Zero-day chosen as $fE_s > 5$ Mc/s.

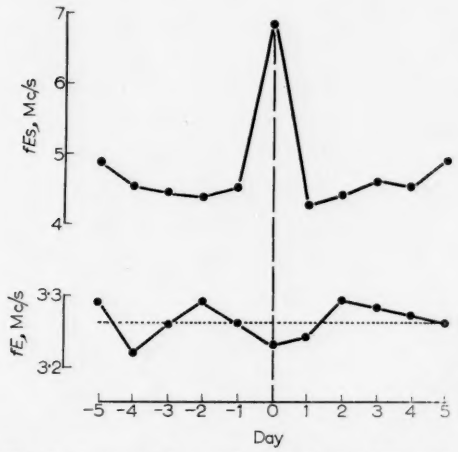


Fig. 2. Superposed–epoch analysis of fE_s and fE data for de Bilt, all hours 1954–1957, for same days as used in Fig. 1.

VOL.
14
1959

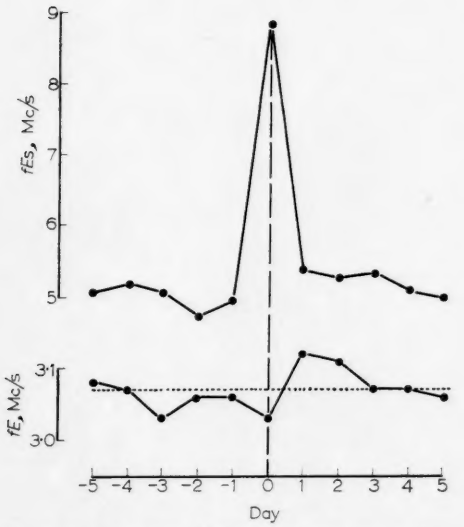


Fig. 3. Superposed–epoch analysis of noon fE_s and fE data for Slough, 1952–1954. Zero-day chosen as $fE_s > 7$ Mc/s.

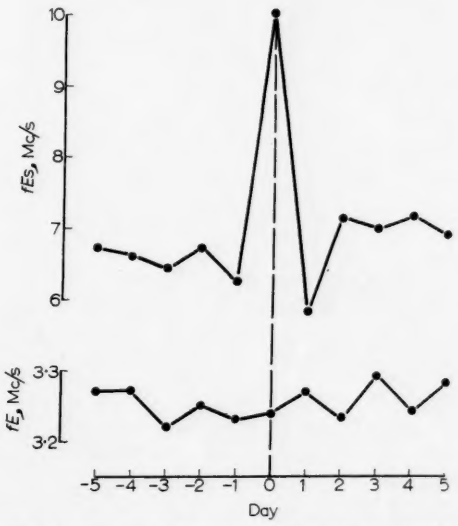


Fig. 4. Superposed–epoch analysis of noon fE_s and fE data for Tokyo, 1952–1954. Zero-day chosen as $fE_s > 9$ Mc/s.

for fE_s does not appear to produce any significant depression in fE from the average value for the 11 days examined.

The conclusion to be drawn from the superposed-epoch analyses for these three stations is that on days of very intense E_s-ionization (fE_s some 4 Mc/s *above average*) a depression of fE of about 0·03 Mc/s at most may occur. It is to be emphasized that in practice the high threshold values of fE_s chosen for zero-day are only exceeded for a comparatively small fraction of the total observing time; e.g., at Slough, on the average, fE_s at noon exceeds 7 Mc/s for only about 15 per

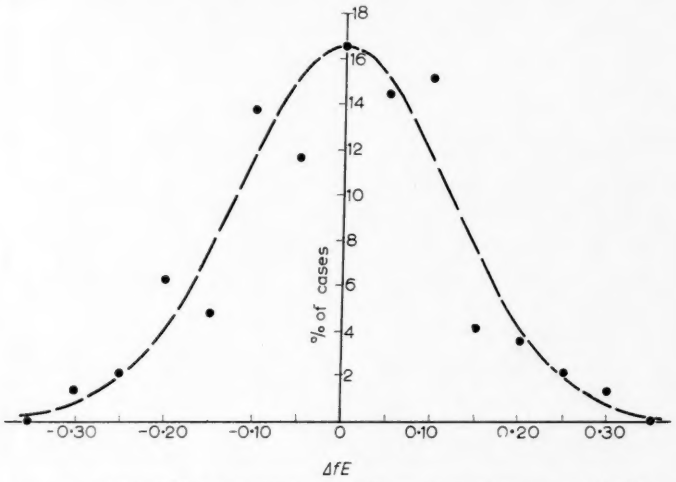


Fig. 5. Distribution of departures of fE on days when $fE_s > 5$ Mc/s at Slough for 1944–1955. Gaussian error curve shown as broken line.

cent of the summer days, and for only about 5 per cent of the winter days. Hence in dealing with monthly mean data, the depressions in fE due to the influence of E_s-ionization are likely to be considerably less than those indicated by the superposed–epoch analysis.

If the order of magnitude of the decrease in $h'E_s$ with intensification of E_s-ionization illustrated in Fig. 1 is typical, then the fact that this only gives rise to a reduction of fE of about 0·03 Mc/s suggests that the level of the E_s-ionization is probably very near to that of the normal E-layer peak. An alternative possibility is that on those occasions when E_s-ionization is significantly below the normal E-layer peak it is still sufficiently transparent to enable penetration of the normal E-layer to be observed.

In a further attempt to detect a possible influence of E_s we have considered the noon values of fE observed at Slough over a period of 12 years on those days for which fE_s was greater than 5 Mc/s. The departures, ΔfE, of these values from a smooth curve drawn through the remaining values of fE (i.e. those for days on which $fE_s < 5$ Mc/s) were determined. If the measured values of fE were spuriously low on the days of enhanced E_s then one would anticipate a preponderance of negative values of ΔfE. Fig. 5 shows the distribution actually obtained for all data over the period 1944–1955 for Slough. It will be seen that in fact positive and

negative values occur with equal frequency and that there is no significant departure from a normal error distribution (shown as broken line).

From all these results, it would appear that there is fairly conclusive evidence that even in the presence of the most intense E_s-ionization it is possible for the scaling of fE to be practically unaffected.

4. Influence of the S_q-Current System

4.1. *Vertical drift*

It has been shown by Martyn (1953) that horizontal movement of an ionized region across the horizontal component of the earth's magnetic field can give rise

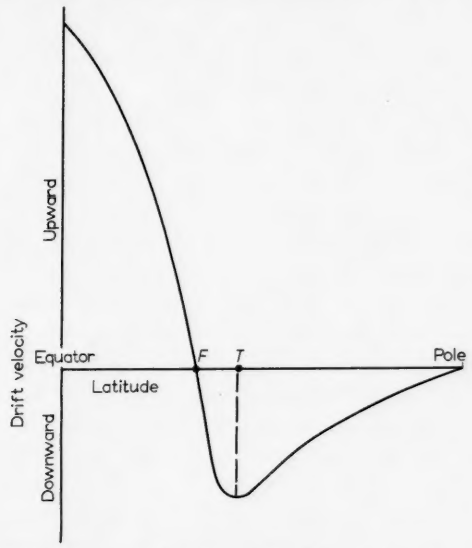


Fig. 6. Diagrammatic representation of the latitude variation of the vertical drift velocity in the E-region, resulting from the influence of the S_q current system.

to quite marked vertical drifts of ionization, resulting in changes in the electron density–height profile of the region. The magnitude of the vertical drift velocity is determined by two factors: (a) the east–west component of the current, and (b) the strength of the horizontal component of the earth's magnetic field (approximately north–south). The variation of drift velocity with latitude has also been discussed by Martyn (1947a), who has shown that it has the form illustrated diagrammatically in Fig. 6. At the latitudes of the foci of the S_q systems (F), the factor (a) is zero, and at the poles factor (b) is zero, so that no vertical drift occurs. Near 1100 hours LT, at latitudes between the foci and the equator, the S_q-current is directed towards the east, so that vertical drift will be directed upward, and since the horizontal component of the field increases towards the equator the velocity will be a maximum at the equator. At the same time the current is towards the west at latitudes greater than that of the foci, and vertical drift is downward, with a maximum value at some intermediate latitude T. At all latitudes, the magnitude of vertical drift will change with local time and with the seasonal

variation in the strength of the currents. Furthermore, the current system becomes distorted in the vicinities of the magnetic equator and the auroral zone, where there appear to be narrow belts of maximum current density (the so-called "electrojets").

As might be expected, the variation of the level of maximum electron density as the result of vertical drift is directly related to the drift velocity (equation (14)). However, according to equation (13) the change of maximum electron density N_m due to the influence of vertical drift depends both on the drift velocity and on the vertical height gradient of this velocity. A rough estimate of the relative magnitudes of the two terms in equation (13) may be gained by considering the distorting influence for a height change $(\Delta h)_{N_m}$ of say 2 km. Insertion in equation (14) of typical values of α and N_m for summer noon at a temperature latitude station gives $v \sim 500$ cm/sec. A reasonable estimate of the velocity gradient $(\partial v/\partial h)$ in the E-layer would appear to be of the order of 10 cm/sec per km (Shimazaki, 1958). Substituting these values in the right-hand side of equation (13) and assuming a scale height of 8 km shows that the term involving the velocity gradient would give rise to a 4 per cent change in N_m, whereas the second term would contribute 1–2 per cent. It is to be noted that the term in v^2 alone depends on the height change considered, and for $(\Delta h)_{N_m} = 1$ km its magnitude is reduced by a factor of 4. This indicates that at middle and high latitudes the *vertical gradient* of drift velocity plays the major part in determining the resultant distortion of the layer. We may reasonably assume that the magnitude of $\partial v/\partial h$ will change directly with that of v (both becoming zero at the foci), but it is not possible to generalize on the type of variation of the vertical velocity with height. At levels above that of maximum S_q-current density the magnitude of the vertical velocity within the current sheet will decrease with increase of height. In this height range, v and $\partial v/\partial h$ will always be of opposite sign, and $\partial v/\partial h$ will be negative on the equator sides of the foci, and positive on the poleward sides. Thus, if the level of the E-layer peak (N_m) is above that of the maximum of the S_q-current, then according to equation (13) at middle and high latitudes downward drift will be associated with a depression of fE, and upward drift with an elevation of fE. It may be noted that Appleton *et al.* (1955) and Shimazaki (1957) have given experimental evidence for a vertical drift influence on fE in this sense. For equatorial latitudes, the influence of the second term of equation (13) may modify the conclusions with regard to the effect of upward drift on N_m, but the magnitude of the velocity of drift would have to be very large for this to be the case.

VOL
14
195

4.2. *The index n*

We have found that a very convenient initial approach in the treatment of values of E-region critical frequency is to use the simplified equation (8), according to which the relation between log cos χ and log fE should be linear. When considering seasonal changes in monthly mean noon values of fE it is necessary to make appropriate allowances for changing solar activity, and when this is done it is found that the logarithmic plots are in fact substantially linear. The index n is then determined by the method of least squares.

In the simple equations (8) and (11), the movement term div $(N\boldsymbol{v})$ of equation (1) has been omitted, and the assumption made that the maximum ionization

density occurs at the height of maximum production. If these omissions are serious we might expect them to influence the values of the parameters n and A determined experimentally.

4.3. *Variation of n with latitude*

We have examined monthly mean noon data for more than forty widely scattered stations ranging from 71°N to 65°S, and find that the index n shows

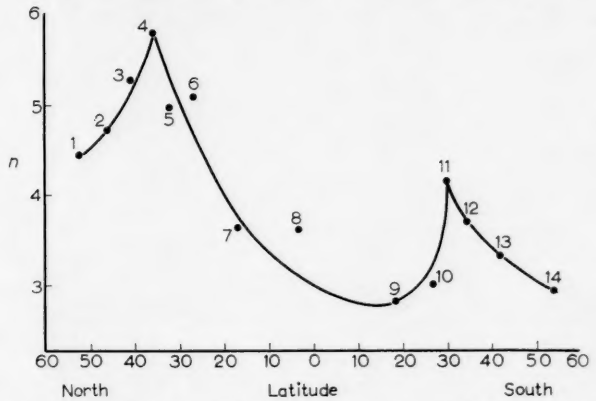


Fig. 7. Index n for stations centred approximately on the 150°E meridian, derived from noon values of fE for 1952–1954. Station key: (1) Adak; (2) Wakkanai; (3) Akita; (4) Tokyo; (5) Yamagawa; (6) Okinawa; (7) Baguio; (8) Singapore; (9) Townsville; (10) Brisbane; (11) Watheroo; (12) Canberra; (13) Hobart; (14) Macquarie Island.

certain systematic variations with latitude. This can be illustrated by considering the latitude variation for the fairly complete network of stations located in the longitude zone centred near 150°E. Values of n (deduced by the method of least squares) for stations in this zone are shown in Fig. 7. In computing n we have used monthly mean noon data for 1952–1954; these are years of low solar activity and only minor corrections have been necessary to allow for varying month-to-month activity. It will be seen that there is a well defined variation with latitude, and although it is true that n is smaller at equatorial than at temperate latitude stations, it attains sharp maximum values at Tokyo (35°N) and Watheroo (30°S). A limited investigation of data for the period 1947–1949 (years of high solar activity) shows that the peak value of n at Tokyo was still present, or even accentuated.

A second striking feature of the curve of Fig. 7 is the marked asymmetry with respect to the two hemispheres. We find that part of this arises from an asymmetric influence on fE. The value of n is a measure of the seasonal variation in noon values and consequently it is very sensitive to any influence which affects fE unequally in summer and winter. Now the solar distance is some 3 per cent smaller in January than July, and this may be expected to introduce slight asymmetry into the fE values. Suitable correction for this gives rise to the curve shown in Fig. 8. It will be seen that although most of the asymmetry is now removed, some still remains in the higher northern and southern latitudes. It seems possible that this residue may be associated partly with the magnetic asymmetry of the earth and partly with some residual sunspot effect.

We have made a similar study of n for stations in two other longitude zones centred on 15°E and 70°W. In these cases the networks of stations are not as complete as for the 150°E meridian, but it is clear that in the northern hemisphere, at least, singularities exist near latitude 30°N. In the case of the southern hemisphere, our results indicate a marked increase in n towards the latitude of Capetown (34°S), and in the case of the South American stations there may be a singularity at some latitude between Buenos Aires (35°S) and Falkland Islands (52°S). The curves for these two longitude zones also exhibit asymmetry between the two hemispheres.

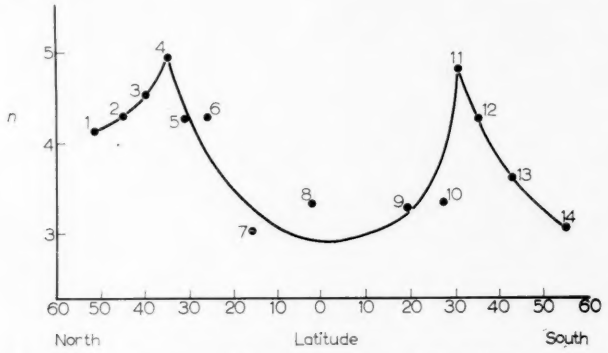


Fig. 8. Index n for stations centred approximately on the 150°E meridian, derived from noon values of fE for 1952–1954 and reduced to a common solar distance. Station key as for Fig. 7.

4.4. *Explanations of the singularities in n*

The high values of n obtained for stations about 30° to 40° north and south of the equator strongly suggest that the values of fE are in some way influenced by the ionospheric S_q-current systems, the foci of which are located at about $\pm 35°$ latitude. Possible explanations of the singularities can, in fact, be formulated in terms of (a) a seasonal shift in the latitudes of the foci of the S_q-current systems; (b) the seasonal change in the intensity of the S_q-current systems; (c) a seasonal change in the vertical gradient of vertical drift velocity.

Before discussing these possibilities it may be mentioned that studies of the worldwide distribution of sporadic E-ionization indicate that this ionization appears to exhibit a latitude variation having certain characteristic features near the latitudes of the S_q-foci. However, the magnitude of these particular E_s-changes is very small, and in view of the results outlined in Section (3.2) we do not consider that the sharp cusps of Fig. 7 can be explained in terms of an influence of sporadic-E on the published values of fE. The general latitude and diurnal variations of E_s-ionization noted above suggest that the sporadic-E layer is also influenced by the S_q-current system. We shall now consider (a), (b) and (c) above.

(a) The remarkable cusp-like appearance of the singularities in Fig. 7 suggests that these may arise from some differential effect on fE over the seasons. Any influence which changes fE equally in summer and winter will not produce a marked variation in the index n, but if fE is reduced in one season and increased in the other then quite a pronounced change in n may be expected. The peak in n

at Tokyo in Fig. 7 could be readily explained if, in December, the effect of the current system is to increase fE and, in June, to reduce it.

The differential effect will thus be produced if we assume that Tokyo is situated at the mean latitude of the S_q-focus, and that at the two solstices the focus moves respectively to the north and south. If the sense of the change of fE with the direction of drift is as postulated in Section (4.1) this suggestion requires that the current focus is to the north of Tokyo in December and to the south of Tokyo in June. We shall discuss below this question of the seasonal movement of the current foci.

(b) An alternative explanation of the maximum values of n in Fig. 7 is to suppose that they arise from the seasonal change in the strength of the current system. The latter is such that the vertical drift effects will be considerably greater in local summer than in winter. If we assume that Tokyo is situated at the latitude of maximum downward drift (indicated by T in Fig. 6), and that such drift is associated with depressions of fE, then the June values of fE will be reduced more than the December values, and the slope n will be abnormally high.

(c) Referring to equation (13) and the remarks in Section (4.1) it is clear that large values of the index n can arise if there are appropriate seasonal variations in the height gradient of vertical drift velocity. In this case, again, maximum values of n would be anticipated at the latitude of maximum downward drift (T in Fig. 6). In this connexion it is worth noting that there is direct experimental evidence for marked seasonal changes in vertical gradients of *horizontal* velocity in region E (Greenhow and Neufeld, 1956; Huxley, 1957), and these might be expected to lead to corresponding changes in gradients of vertical velocity.

VOL. 14 1959

From the foregoing it is apparent that on the basis of any of the above interpretations, these ionospheric data indicate that the mean latitudes of the S_q-foci along the 150°E meridian are respectively equal to, or less than, 35°N and 30°S. It is interesting to note that an analysis of the *magnetic* data provided by Vestine *et al.* (1947) indicates that, for the 150°E longitude zone, the average positions of the S_q-foci may be placed within the ranges 30–35°N, 25–30°S. These averages are based on data for the International Polar Year (1932–1933) and for the period 1922—1933, and refer to periods of low sunspot activity. The fact that the figures for the average latitudes of the foci north and south of the equator obtained from ionospheric and magnetic studies agree may be taken as strong support that the interpretation of the ionospheric results in terms of the influence of the S_q-current system is substantially correct.

World maps of the S_q-current systems have been given by Chapman and Bartels (1940) for a year of sunspot minimum. These indicate a shift of about 10° in the latitudes of the foci towards the equator in local summer, i.e. in the sense required by explanation (a) above, and in an earlier preliminary report on this work (Beynon and Brown, 1956) the authors inclined to explain the results in this way. A seasonal shift of the foci in the same sense has also recently been reported by Endo (1957). On the other hand, an analysis of the very comprehensive data on the variations in the terrestrial magnetic field published by Vestine *et al.* (1947) shows that over the 12 year period 1922–1933 the average positions of the S_q-current foci, considering stations distributed over the world,

were slightly nearer the equator in local winter than summer. However, a closer study of these more recent magnetic data indicates that the seasonal shift varies both with respect to the magnitude and sense of the shift for stations situated along different meridians of longitude. Furthermore, evidence for considerable change in the location of the foci, even over successive days, has been given by Hasegawa and Ota (1937). The experimental evidence at present available concerning the sense of the seasonal shift of the foci is thus inconclusive.

Recently, we have obtained some magnetic data for the Japanese observatories Kakioka (36°N) and Memambetsu (44°N) for the period of the ionospheric observations discussed here (1952–1954). We have examined the diurnal variation in the

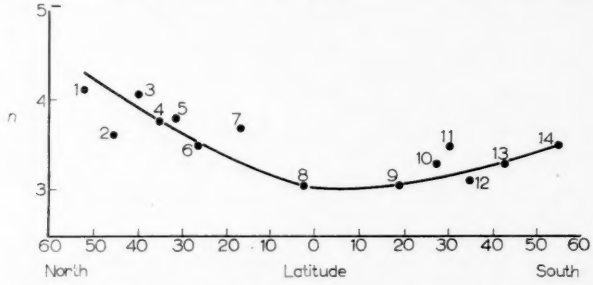


Fig. 9. Index n for stations centred approximately on the 150°E meridian, derived from values of fE at 1500 hours for 1952–1954. Station key as for Fig. 7.

horizontal component of the earth's magnetic field at these stations, and find that for each of the years 1952–1954 the form of the variation at both places in both June and December is that to be expected for stations situated *north* of the S_q-focus. Hence it would appear that for the years considered and for the 150°E longitude zone the latitude of the focus of the northern hemisphere S_q-current system was always less than 36°N, i.e. to the south of Tokyo, in all seasons. This would appear to rule out an explanation of the peaks in terms of (a) alone.

The possibilities (b) and (c) remain, and it seems likely that the actual latitudes at which the peaks in n occur are determined by the net influences of (a), (b) and (c). There is little doubt that seasonal variations in the latitudes of the current foci do occur and that such variations will influence but not solely determine the latitudes of the peaks. Supporting evidence for a seasonal variation in the influence of vertical drift in the sense postulated in (b) will be given in Section (4.9).

Mention might also be made of a further possibility. If the vertical depth of the S_q-current sheet is small, seasonal changes in the height of the layer peak, and/or of the current sheet, may result in a seasonal variation in the magnitude of the resulting geomagnetic distortion. Evidence for distinct stratifications even within the E-region has been obtained in recent rocket experiments (see Section 4.8.2. below), and it seems possible that the E-layer peak may sometimes be in a current sheet and at other times above or below it. The observed marked reduction in the seasonal variation of fE (i.e. large values of the index n) at certain latitudes may thus possibly be explained in terms of the seasonal variations in the relative levels of the current sheet and of the E-layer peak. However, with present knowledge it is not possible to pursue this suggestion further.

4.5. *Condition of no vertical drift*

Additional evidence for the influence of the S_q-current system on fE is provided by a study of the latitude variation of n for 1500 hours. Near this time, at all latitudes less than about 70°, the east–west component is very small so that vertical drift effects should also be much reduced. Fig. 9 shows n at 1500 hours for the same stations considered in Fig. 7. It will be seen that at this hour there are no significant singularities near latitudes 35°N or 30°S. If the values of fE are adjusted for the varying solar distance over the year, the asymmetry of the curve in Fig. 9 is again removed, giving a minimum value of $n \sim 3$ at the equator.

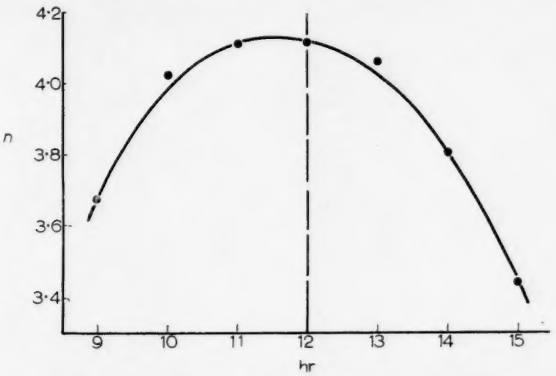


Fig. 10. Diurnal variation of index n derived from values of fE for Slough for 1952–1954, reduced to the common sunspot number $R = 16$.

When the influence of the current systems is absent there remains a gradual decrease in n towards the equator and this may represent a real change in E-layer characteristics (such as vertical gradient of temperature or recombination coefficient) with change of latitude (see equation (11)).

4.6. *Variation with longitude*

It is interesting to note that the peaks in n found for stations situated approximately on the 75°W meridian (see Section (4.3)) suggest that at this longitude the S_q-foci are situated in latitudes $\leqslant$30°N and $\leqslant$(40–50°)S, i.e. some 10° southward compared with those for the 150°E meridian. Analyses of magnetic data for the 75°W longitude zone by Vestine *et al.* (1947) and McNish (1937) also show that the mean positions of the S_q-foci follow the 10° southward displacement of the magnetic equator.

4.7. *Diurnal variation of n*

At any given station the east–west component of the S_q-system is a maximum on the average at about 1100 hours. Hence it would be expected that any ionospheric influence of the system would show an asymmetric variation with respect to local noon. This prediction is confirmed by Fig. 10 which shows the index n for Slough (51·5°N) determined as before by the method of least squares for each hour between 0900 and 1500 hours. It is seen that n varies asymmetrically over the day, forenoon values always being greater than the corresponding afternoon ones, so

VOL. 14 1959

that the maximum is displaced about 1 hr before noon. A similar analysis for Tokyo yields the same result, but the diurnal change in the value of n for Tokyo is appreciably larger than that at Slough as might be expected.

4.8. *Diurnal variation of fE*

4.8.1. *Temperate latitudes.* Evidence for a detectable depression in E-layer electron density in summer near 1000 hours at a north temperate station which can be ascribed to the influence of the dynamo current system has been given by Appleton *et al.* (1955). In order to examine whether the magnitude of this

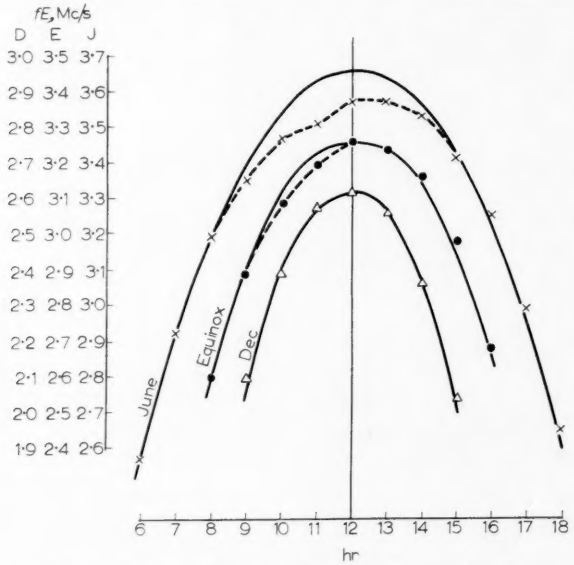


Fig. 11. Diurnal variation of fE at Slough for solstice and mean equinox months, 1944–1953. Full-line curves: theoretical variation; × observed June; ● observed equinox; Δ observed December.

depression varies with the seasonal change in the strength of the S_q-currents we have considered the diurnal variation of fE for the solstice and equinox months for Slough, averaged over 10 years, and the results are shown in Fig. 11. The curves have been fitted to the observations for the early morning and late afternoon hours, and they represent the theoretical variation of fE to be expected in the absence of any perturbing influences. It will be seen that in summer significant depressions in fE occur over the period 0900–1400 hours with a maximum depression near 1100 hours. At the equinox the perturbations are much smaller, and they are apparently absent in winter. These long-term averages would thus appear to confirm that downward drift is associated with a decrease in ionization density, and that the effect is most marked when the current is a maximum.

However, a study of the variation in fE on individual days shows that a depression certainly does not occur at 1000–1100 hours *every* day. The experimental evidence is that on most days a significant "bite-out" appears at some time between, say, 0900 and 1400 hours, and examples can be found when the diurnal variation appears

quite free from any perturbing influence. It is also worth noting that whereas a bite-out can often be detected on most days within a month, it is sometimes found that this is smoothed out in the monthly mean diurnal curve because the hour of occurrence varies so much from day to day.

If the above-mentioned depressions in fE are attributed to the influence of the S_q-current system, then the variability in the time of occurrence of the depression is not unexpected, since it is well established that there is a considerable day-to-day

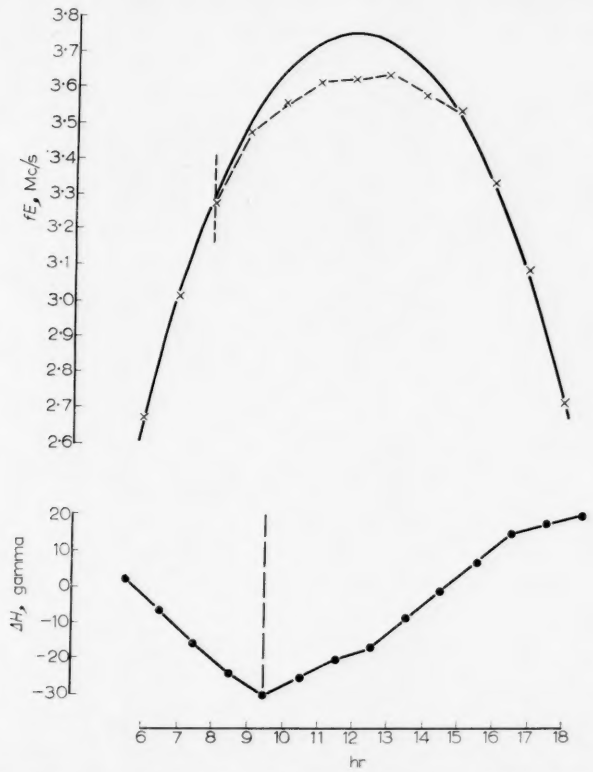


Fig. 12(a). *Lower curve*: diurnal variation of ΔH (departure of the horizontal component of the magnetic field from the daily mean value) at Abinger for June, 1955–1957, averaged for all days when the minimum value occurred between 0900 and 1000 hours GMT. *Upper curve*: diurnal variation of fE at Slough for June 1955–1957, averaged for the same days used in the lower curve. Full-line curve: theoretical variation; × observed values.

variability in the position of the S_q-current system relative to the noon meridian and also in the intensity of the currents (BARTELS, 1949). On those days when the S_q-currents exert no perturbing influence and the diurnal variation of fE is quite smooth, it may be that the currents are weak or that the level at which they flow is not sufficiently near the E-layer peak (see Section (4.8.2) below.) The day-to-day variability in the magnetic data can, in fact, be usefully employed to demonstrate the associated distortion in the E-layer. For this purpose we consider data for two groups of days in the month of June for the 3 years 1955–1957. The first

group consists of all those days (34 in number) on which the maximum east–west component of the S_q-current system occurred at this latitude between 0900 and 1000 hours GMT, and the second all days (18) for which the maximum occurred at a significantly later time, viz. between 1200 and 1300 hours GMT. The time of maximum east–west current was obtained from hourly values of the horizontal component of the earth's magnetic field observed at the Abinger Magnetic Observatory. Figs. 12(a) and 12(b) show the mean departures of the horizontal

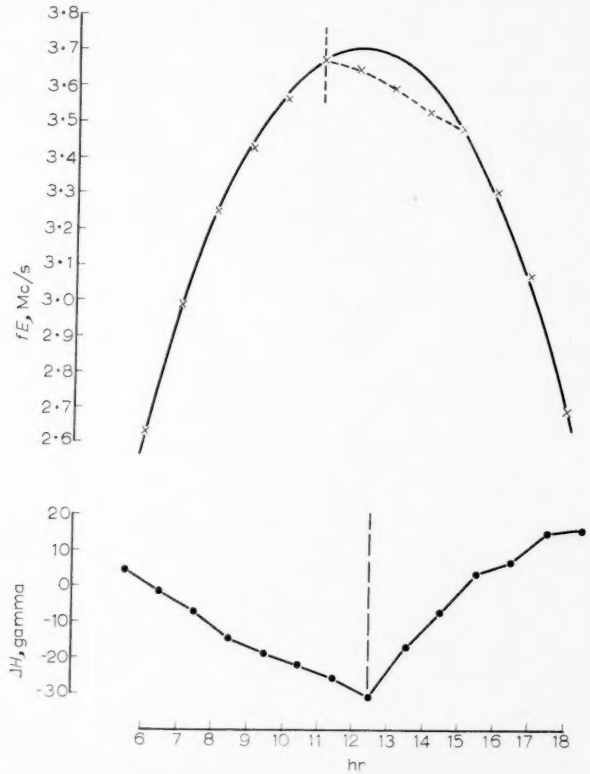


Fig. 12(b). *Lower curve*: diurnal variation of ΔH (departure of the horizontal component of the magnetic field from the daily mean value) at Abinger for June 1955–1957, averaged for all days when the minimum value occurred between 1200 and 1300 hours GMT.
Upper curve: diurnal variation of fE at Slough for June, 1955–1957, averaged for the same days used in the lower curve. Full-line curve: theoretical variation; × observed values.

component of the field between the hours 0600 and 1800 and the corresponding mean diurnal variation in the critical frequency. The curves give the theoretical variation of fE fitted to the early morning and late afternoon values. It will be seen that there is a perceptible difference in the behaviour of fE in the two groups of days. In the first group, definite departures are to be seen from 0900–1000 hours onwards, whereas in the second group no significant effect is noticed until 1200–1300 hours. In both groups the recovery seems to be complete by 1500 hours as might be expected. It would appear too that the magnitude of the depression in

fE is larger for the first group than for the second. This may be associated with the fact that in the morning the vertical drift due to the S_q-currents is in the same sense as the normal diurnal change in the level of the layer, whereas in the afternoon the height changes are in opposite senses.

It is relevant to note that the mean diurnal variation of fE_s for the two groups of days showed *no* significant differences, thus confirming that the fE-changes exhibited in Fig. 12 are not to be attributed to an influence of sporadic-E ionization.

In view of the possible ambiguity involved in scaling the parameter fE, outlined in Section (3.1), we have also examined a number of *individual* ionograms taken at stations in this latitude, and carefully measured the diurnal variation in the parameters indicated by fB and fA in Fig. 13. These types of transition from

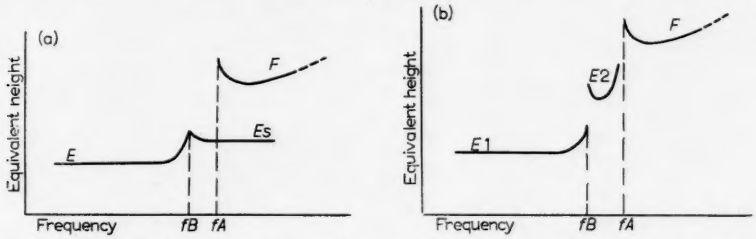


Fig. 13. Typical $h'f$ traces observed in the transition from the E- to F-layers.

the E- to the F-layer are of comparatively frequent occurrence, and although the true value of fE is not directly observable one can infer that it must lie between fB and fA. In fact, there are good grounds for believing that in Figs. 13(a) and (b) the parameters fA and fB respectively are fairly reliable measures of the true E-layer critical frequencies. It is also to be emphasized that although the magnitude of the frequency fB may be affected by the presence of an E_s-trace (cf. Section (3.2)), the value of the "threshold" frequency fA will not be so affected; provided, of course, that the cusp which it denotes remains.

Fig.14 shows a plot of fA and fB carefully scaled from ionograms taken at $\frac{1}{4}$ hr intervals. Points have only been plotted where the characteristics sought were clearly defined on the records. It will be seen that from 0600 to 0900 hours there is no significant difference between fB and fA, but over the noon period a frequency gap developed, and between 1000 and 1200 hours *both* parameters show a distinct depression. This behaviour is typical of many records which we have examined. The fact that such pre-noon depressions are not confined to the parameter fB may be taken as strong evidence that the underlying cause does not result from a mere error in scaling, but really represents a true reduction in the peak electron density of the E-layer at this time.

In the preceding paragraphs we have discussed the S_q-distortion of the E-layer in terms of a simple change in level together with an associated perturbation in the peak electron density. This picture is, of course, an oversimplification and in practice the type of distortion produced might be expected to be a more complicated phenomenon probably involving a stratification of the layer determined by various factors, and in particular by the height variation in the current density. The examples just quoted, in which both fA and Bf exhibited a pre-noon depression,

suggest that the geomagnetic distortion was here effective over an appreciable height range (see Section (4.8.2) below). In this connexion it is interesting to note that recent direct observation of the current system by rockets (Cahill and Van Allen, 1958) shows that a stratification in the vertical distribution of current density does in fact occur.

4.8.2. *Equatorial latitudes.* If upward drift is to be associated with an elevation of fE, then we might expect to detect a corresponding elevation near 1100 hours

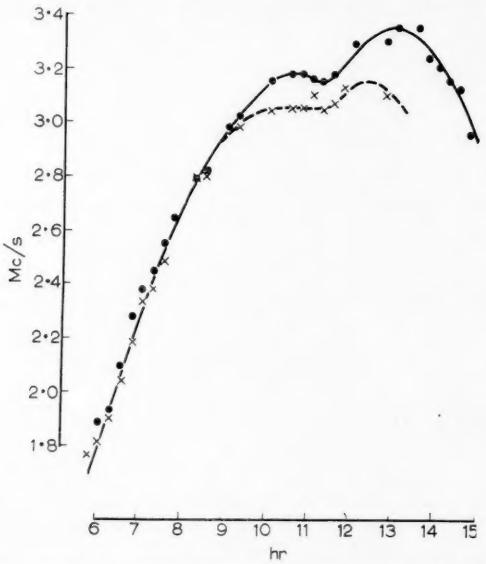


Fig. 14. Diurnal variation of parameters $fB(\times)$ and $fA(\bullet)$ for Slough, 21 September 1951, obtained from ionograms taken at $\frac{1}{4}$ hr intervals.

in the diurnal variation for equatorial stations. As yet, we have not found any conclusive evidence for this. Isolated examples can be found in which the monthly mean diurnal variation shows a slight pre-noon elevation but an examination of the mean data for Singapore (1°N), for example, shows that, if anything, there is at this station a slight reduction near 1100 hours. It should be added that for Singapore we have only been able to use limited data compared with the 10 year averages used for Slough and Tokyo, and this makes the detection of small departures from "normality" difficult. We have also examined data for Huancayo (12°S), a station at which the amplitude of the S_q-diurnal variation is known to be exceptionally large, and Fig. 15 illustrates the mean diurnal variation of fE, averaged over 4 years, for a summer month (February) and for the winter (June). As in Fig. 11 the curves have been fitted to the observations for the early morning and late afternoon hours. It is clear that, whereas there is a suggestion for a small elevation of fE over the hours 1100 to 1300 in June, no significant departure from the theoretical variation is apparent in the month of February. Examination of data for other months shows that over the noon period departures from the theoretical curves are in all cases very small and as likely to be negative as positive. A similar conclusion is obtained for the diurnal variation of fE at Ibadan (7·4°N),

VOL
14
195

another place at which the daily range of the horizontal component of the magnetic field is very large. We may therefore conclude that in equatorial latitudes there is little evidence for layer distortion in the *diurnal* variation of fE.

It is to be added that a study of the *latitude* variation in fE certainly indicates an equatorial enhancement of the peak electron density (APPLETON, 1955; SHIMAZAKI, 1958). This is illustrated in Fig. 16 which shows the variation with solar zenith angle χ of fE at noon for temperate and equatorial stations corrected to a

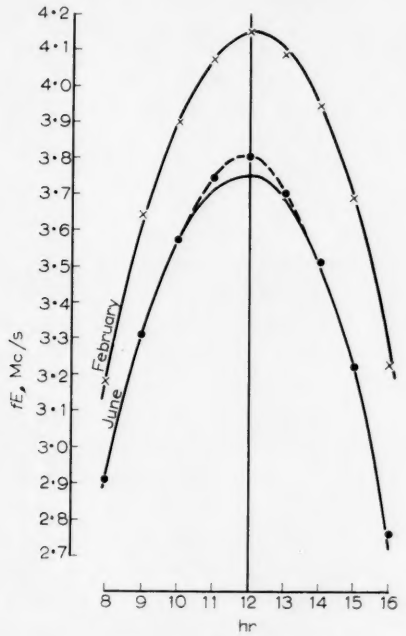


Fig. 15. Diurnal variation of fE at Huancayo for summer (February) and winter (June) months, 1938–1941. Full-line curves: theoretical variation; × observed February; ● observed June.

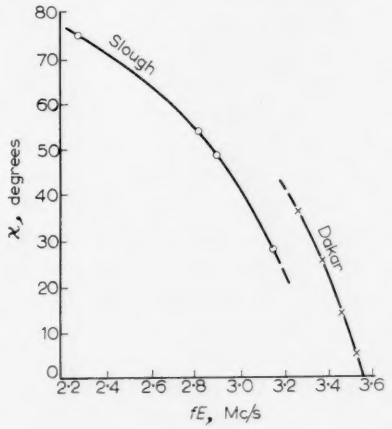


Fig. 16. Variation of noon values of fE with solar zenith angle χ for Slough (latitude 51·5°N) and Dakar (latitude 14·7°N). The values of fE correspond to zero sunspot number.

common sunspot number. It will be seen that there is a clear discontinuity between the curves and that over the range of χ which is common to both latitudes the equatorial values of fE are larger by about 0·2 Mc/s than those for temperate latitudes. These curves have been carefully determined from a large body of data, and the magnitude of the gap between them greatly exceeds the uncertainty of the points. Part but not all of this gap must be attributed to the fact that the Slough values in summer are depressed owing to S_q-effects. The residue could be attributed to the influence of upward drift at Dakar in elevating fE. However, it is difficult to reconcile this simple explanation with factors such as the absence of any consistent enhancement in the diurnal variation of fE during the pre-noon period, and the fact that equatorial values of fE also seem to be elevated with respect to temperate latitudes at hours when vertical drift effects would normally be expected to be small.

Further consideration of this apparent "equatorial anomaly" must await additional data, but it should be remembered that the net distortion of the E-layer resulting from a vertical drift velocity depends on several factors. In equation (13), which gives the net distortion on the assumption that both H and α do not vary over the vertical height concerned, the velocity gradient term in general dominates, as we have seen. However, at equatorial latitudes the term involving the velocity itself will assume increasing importance (Fig. 6). Furthermore, if there were an appreciable positive height gradient λ of scale height H at equatorial latitudes, an elevation of the height of the E-peak would result in a lowering of fE, in opposition to the effect of the height gradient of velocity. In fact, it may be noted that the latitude change in n observed in the absence of vertical drift (Fig. 9) might be taken as evidence for a positive value of λ at the equator. In addition, there is a possibility that at the equator the peak of the E-layer is situated such that the distorting influence of the S_q-current sheet is not large. In this connexion it is relevant to note the experimental results obtained by Singer *et al.* (1952) and by Cahill and Van Allen (1958) with rocket-borne magnetometers at the magnetic equator. The earlier results indicated that the equatorial electrojet was confined to a narrow range of altitude in the *lower* portion of the E-layer (93–105 km). On the other hand, the results from the more recent rocket flight provide evidence for at least *two* current sheets, the lowest extending from 97 to 110 km, and the second from 118 km upwards. This experiment also suggested that only about one-half of the total current system lay between 97 and 121 km. Between 110 and 118 km the measurements indicated little or no current density, and it is clear that if, at the time concerned, the level of the E-layer peak lay within this height range then there would be no S_q-distortion detectable in fE.

It seems reasonable to anticipate that the stratification in the current density already observed at the equator exists at other latitudes. It is probable too that the comparatively large vertical depth of the current system found in the equatorial experiment also occurs at other latitudes. If this is so, then it might explain a number of features of the E-trace on ionograms. Thus the apparent absence of any S_q-distortion in fE on certain days can be understood, as well as its absence on days when distortion can be detected in $h'E$ and/or for levels above the layer peak. Some of the stratifications observed on ionograms in the E-trace and during the transition from the E- to the F-layer might also be associated with a stratified current system.

4.9. *Magnitude of the S_q-effect on fE*

An approximate estimate of the magnitude of the drift effect on fE at a temperate latitude station can be obtained from the curves of Fig. 11. These show that the maximum depression in fE is about 0·12 Mc/s, or 3–4 per cent, and occurs, as expected, at 1100 hours in the summer. At noon, the depression would appear to be about 0·09 Mc/s (2–3 per cent) in June. For the equinox months, the maximum depression is of the order of 0·03 Mc/s, while no measurable effect can be detected in December. These figures are based on long-term monthly mean data for Slough and represent average values for a mean sunspot number of 76.

For Tokyo, an analysis of data covering the period 1946–1956 (mean sunspot

number = 82) shows a seasonal change in the magnitude of the depression in *fE* near 1100–1200 hours in the same sense, but the changes are all larger than for the case of Slough. In June, the depression in *fE* at Tokyo is about 0·14 Mc/s (4 per cent) at noon, about 0·08 Mc/s at the equinoxes, and about 0·03 Mc/s (1 per cent) in December.

In an earlier note (BEYNON and BROWN, 1956) we interpreted the singularities in Fig. 7 in terms of a seasonal shift in the latitudes of the S_q-foci as seen in Section (4.4(a)) and on the basis of this interpretation we made rough estimates of the magnitude of the changes in *fE* as a result of vertical drift. The estimates now given, although based on a different interpretation, are of the same order of magnitude for temperate latitudes as those given formerly. Our original estimate of the magnitude of the equatorial enhancement of *fE* (5 per cent) was based on the assumption that the abnormally large values at the equator (illustrated for example in Fig. 16) were *entirely* due to the S_q-current. However, the further discussion given in Section (4.8.2) now places difficulties in the way of this interpretation, and the associated estimate of the magnitude must therefore be questioned.

4.10. *Vertical drift and h'E*

VOL. 14 1959

According to the simple Chapman theory, the height of maximum ion production varies with the solar zenith angle (equation (3)) and this changing level manifests itself in the variation of the minimum equivalent height of reflection *h'E* over the day and over the seasons. At the equinoxes, one would therefore expect the noon values of *h'E* to show a latitude variation with a minimum at the equator. However, an examination of the published *h'E* data for many ionospheric stations (MAEDA, 1955; SATO, 1955) shows that *h'E* is substantially independent of latitude. It must be emphasized that the data available are of very limited accuracy, since *h'E* is normally recorded only to the nearest 5 or 10 km, and this might make it impossible to detect height changes of the order of a few kilometres. However, we find that even when data are averaged for many years there is still no evidence for the latitude variation predicted by the simple theory.

An approximate calculation based on equations (3) and (14) indicates that the order of magnitude of this latitude anomaly in the variation of the height of the *E*-layer is such as could be accounted for in terms of the influence of vertical drift resulting from the S_q-current system, in opposite senses at latitudes on either side of the foci. Vertical drift velocities of the order of 500 cm/sec seem adequate to explain the anomaly.

Evidence for an anomalous depression before noon in the *diurnal* variation of *h'E* at a temperate latitude station due to the influence of the S_q-current system has been given by APPLETON *et al.* (1955). It may be noted too that the *seasonal* changes in noon values of *h'E* have been studied at Cambridge (52·2°N) by FINDLAY (1951), who found that although there is agreement with the variation expected for a Chapman layer over the winter period September–March the observed values in the summer months are too low. This might be interpreted as evidence for the influence of westward dynamo currents at a temperate latitude station during the summer.

5. Analysis for Magnetically Disturbed Conditions

5.1. *Magnetic activity and fE*

Previous workers have pointed out that although the major ionospheric effects associated with a magnetic storm occur in region-*F2*, measurable disturbances also occur in the lower layers-*E* and -*F1*, and it is relevant to the present discussion of geomagnetic distortion in region-*E* to consider the behaviour of *fE* under disturbed conditions.

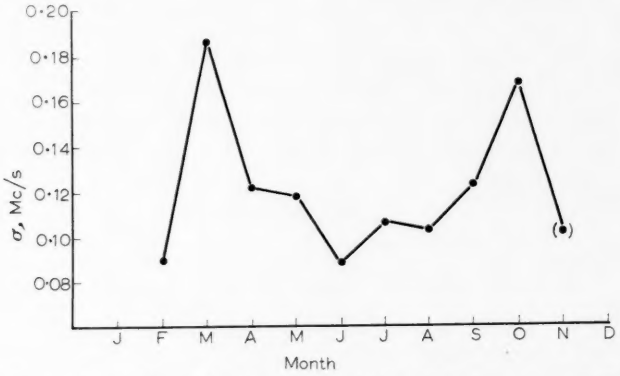


Fig. 17. Seasonal variation of the standard deviation of the departures of *fE* from the mean seasonal curve for 1950–1954 at Kiruna (68°N).

Although the perturbations in *fE* due to magnetic activity are in general small, they are readily detected at periods of high activity (such as the equinoxes) and especially at high-latitude stations (Appleton *et al.*, 1937). Fig. 17 shows the seasonal variation in the standard deviation of *fE* for Kiruna (67·8°N). Although the data for the winter months are rather scanty, there is clear indication for a greater variability in *fE* in the months of enhanced magnetic activity.

At lower-latitude stations we find that it is difficult to establish this type of seasonal variation in the variability of *fE*, since in any month the daily values are often influenced by factors such as the 27 day recurrence tendency associated with the solar rotation period. However, it might be expected that long-period monthly averages of *fE* would show evidence for abnormality at the equinoxes, and we have examined this possibility for monthly noon data for Slough averaged over the 23 year period 1931–1953. Fig. 18 shows a plot of log cos χ against log *fE* for the twelve-monthly mean values, and it will be seen that the points cluster closely to a line, with the exception of those for September and October for which the average *fE* values are too small. It should be added that the data plotted in Fig. 18 inherently contain the quiet-day distortion of the *E*-layer, in that the values of *fE* are progressively depressed as one proceeds from the December to the June point on the line, and this results in a slightly greater slope than would otherwise be the case.

It is tempting to associate the depressions in September and October with increased magnetic activity during these months, and in fact over the years considered the average international magnetic *C*-figures for these months are significantly greater than average. On the other hand, a second maximum of magnetic

activity occurs in March, whereas in Fig. 18 the March point closely fits the line. Furthermore, from work to be described below we find that the magnitudes of the depressions in fE in September and October are several times larger than would be expected from the enhancement of magnetic activity during these months. It might be suggested that this anomaly in September and October may be spurious, but this seems most unlikely in data averaged over 23 years. We have also found that this anomaly occurs in data averaged for hours other than noon. It is also

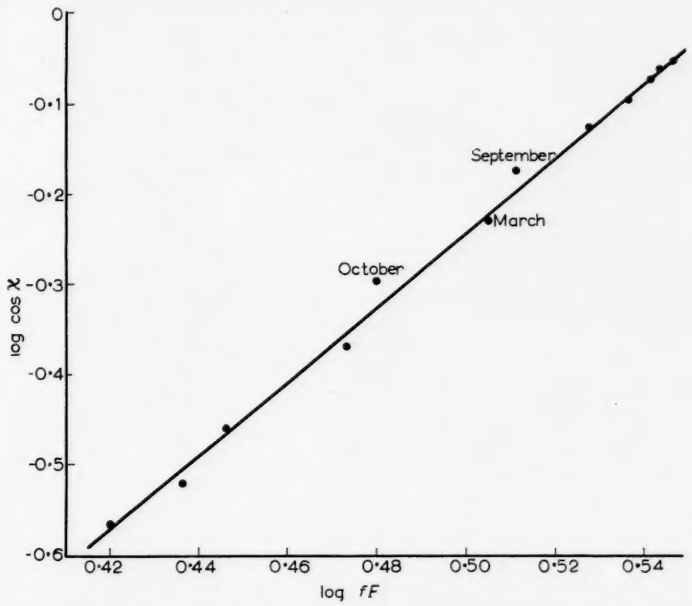


Fig. 18. Seasonal variation of noon values of fE for Slough. Data averaged over the period 1931–1953, and reduced to the common sunspot number $R = 61$.

worth noting that consistently low values of fE in the month of October were observed by WALDMEIER (1944) from a study of 4 years' data at Kochel (47.7°N). It would appear that these depressions in fE observed at some stations during the autumnal equinox are probably real, but cannot be directly accounted for in terms of magnetic activity. It is probable that this anomaly is not unrelated to another "October effect" which has been reported by GREENHOW and NEUFELD (1956).

Another approach to the problem is to apply the superposed–epoch method to the magnetic and ionospheric data. Fig. 19 shows the results obtained from a superposed–epoch analysis of noon fE data for Slough for the years 1950–1952. In this analysis the "zero day" was chosen on the basis of a high value of international magnetic character figure C, and the corresponding fE data were superposed over the period ± 6 days centred on this day. The upper curve shows the mean magnitude of the magnetic disturbance and it is clear that in association with this there are small but perceptible changes in the values of fE. On day -2 there is a small increase in fE, while there is an appreciable depression on day 0, and a minimum value on day $+1$, followed by a slow recovery over 3–4 days. The small

VOL. 14 1959

"positive phase" may possibly be associated with the fact that sunspot areas attain a maximum value some 2–3 days before the date of maximum magnetic disturbance (Stagg, 1928).

For temperate and equatorial latitudes the net disturbance current system is directed westwards at noon, and *if* these currents flow in the ionosphere they might be expected to enhance the effect of the existing S_q-current system for latitudes on the poleward side of the focus, and oppose it for latitudes on the

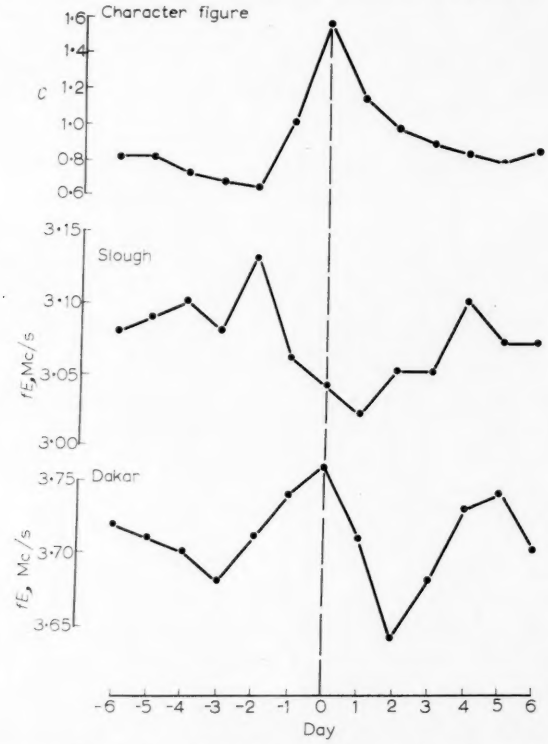


Fig. 19. Superposed–epoch analysis of magnetic and ionospheric data for the period 1950–1952. *Upper curve*: international magnetic character figure; *centre curve*: *fE* at Slough (latitude 51·5°N); *lower curve*: *fE* at Dakar (latitude 14·7°N).

equator side. The superposition of the two systems might thus result in different magnetic disturbance effects in the *E*-layer at temperate and equatorial latitudes.* Thus, according to our discussion in Section (4), at the latitude of Slough during times of magnetic disturbance the quiet-day depressions of *fE* should be enhanced, as is, in fact, the case.

The curves of Fig. 19 indicate that the average depression of *fE* for Slough is 0·07 Mc/s for an increase of the magnetic character figure *C* of 0·7. From this we may conclude that for Slough the depression in *fE* at times of magnetic disturbance is of magnitude comparable with that produced by the S_q-current system (Section (4.9)).

* A similar suggestion has recently been made by Sato (1957).

We have attempted a similar superposed–epoch analysis for an equatorial station, and Fig. 19 shows the result for Dakar (14·7°N) for the same period. In this case the data are far less complete than for Slough, but there appears to be an increase in fE over a period of 1–2 days prior to zero day, followed by a depression to a minimum on day +2. An analysis of considerably more equatorial data is desirable before a definite significance can be attached to this apparently different behaviour between equatorial and temperate latitude stations.

5.2. *The index n on magnetically disturbed days*

We have repeated the calculation of the index n as indicated in Section (4.3) for the four Japanese stations centred on the latitude of the Tokyo singularity, using

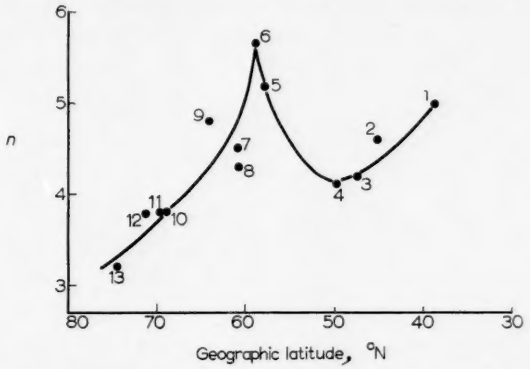


Fig. 20. Index n for stations near and within the northern auroral zone plotted against geographic latitude. Data derived from noon values of fE for 1952–1954. Station key: (1) Washington; (2) Ottawa; (3) St. John's; (4) Winnipeg; (5) Fort Chimo; (6) Churchill; (7) Narsarssuak; (8) Anchorage; (9) Baker Lake; (10) Godhavn; (11) Tromso; (12) Point Barrow; (13) Resolute Bay.

data for only the five international quiet days per month and also the five international disturbed days, over the same period 1952–1954. The results show that the peak value of n at Tokyo is present on both quiet and disturbed days, and that its magnitude is definitely accentuated on disturbed days.

These results again can be interpreted in terms of the enhanced depressions of fE which would occur as a result of downward vertical drift associated with a westward disturbance current system.

5.3. *The index n in the auroral zone*

Fig. 20 shows the results of an extension of the analysis of the index n to northern stations across and within the auroral zone. The data available for high latitude stations are rather limited, and due to the special difficulties of measuring fE at polar stations the results must be treated with reserve, but there would appear to be evidence for large values of the index n in the latitude zone 55–65°N. The largest values are found for Churchill and Fort Chimo, both of which are located on the zone of maximum auroral frequency, while the values for Baker Lake and Narsarssuak, which are close to the auroral zone, are also high.

The value of n is a measure of the sensitivity of the noon values of fE to solar

altitude over the seasons, and the high values found may be interpreted as evidence for a reduced solar control at these auroral stations. The influence of vertical drift produced by the intense currents which flow in the auroral zone is also likely to be considerable. From an analysis of the diurnal variation of fE at Canadian stations, Scott (1952) similarly concluded that the sensitivity of the E-layer to solar angle in the auroral zone is sharply reduced, but that it recovers again to the north of the zone.

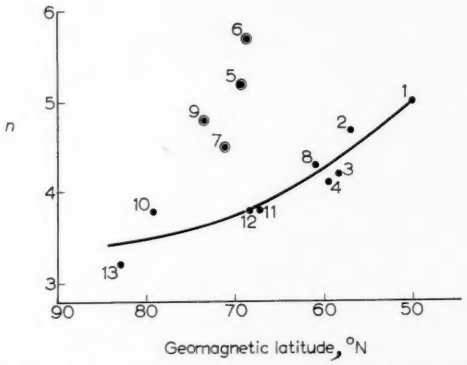


Fig. 21. Index n for stations near and within the northern auroral zone plotted against geomagnetic latitude. Data derived from noon values of fE for 1952–1954. Station key as for Fig. 20.

It is interesting to note that normal values of n are found for Tromso which is also situated on the auroral zone, and for Point Barrow which is close to the zone. The former result is in agreement with the finding of Appleton *et al.* (1937) that the behaviour of the E-layer at Tromso during the Second International Polar Year exhibited the same seasonal control as at a temperate latitude station. Fig. 21 is a re-plot of the points of Fig. 20 on geomagnetic latitude, and from this it would appear that Churchill, Chimo, Baker Lake and Narsarssuak (represented by ringed points) behave differently from Tromso and Point Barrow. The reason for this difference is not clear and further discussion of the phenomenon must be deferred until more extended and accurate experimental data are available.

6. Summary

The present study of the normal E-layer has revealed anomalies which may be attributed to the influence of vertical drift resulting from the S_q-current system. The principal conclusions from this work are:

(1) At a temperate latitude station the distortion is evident as a depression in fE, the magnitude of which is found to vary with season and with local time in the manner to be expected from the established variations of the S_q-current system. For Slough (51·5°N) the maximum distortion (at 1000–1100 hours in summer) has been estimated to be about 7 per cent in peak electron density for a mean sunspot number of 76. The small magnitude of the effect together with difficulties associated with the experimental measurement of fE often tends to mask it on individual days. However, strong evidence for the reality of the phenomenon has been provided

by studying data for different days grouped according to the time at which the distorting influence is a maximum.

(2) At equatorial latitudes the simple drift theory would predict an elevation of fE. A study of the latitude variation in fE makes it clear that there is an equatorial enhancement, but there are difficulties in attributing this entirely to vertical drift, such as the apparent absence of any anomaly in the diurnal variation.

(3) A most striking consequence of the dynamo-perturbation of fE appears from our study of the seasonal variation in noon values at stations over a wide range of latitude, in that the sensitivity to seasonal changes appears to be markedly reduced at latitudes near the foci of the S_q-systems. Possible interpretations of this in terms of drift theory are discussed.

(4) The detection of these small perturbations in fE is, of course, conditioned by the reliability of the parameter fE, and as part of our analysis we have carefully considered the possibility of spurious effects being introduced by the presence of E_s-ionization, but conclude that this is not a major cause of the anomalies considered. A close study of some individual ionograms supports our interpretation of the observed perturbations in fE.

(5) The absence of experimental data on the true height of region-E, and the very limited accuracy of published values of the minimum equivalent height $h'E$ precludes any accurate analysis of these data for height changes which might be attributed to dynamo current influences. However, the absence of any significant latitude variation in $h'E$ and the abnormally low values observed at a north temperate station in summer may result from such influences.

(6) A limited investigation of the perturbation in fE under magnetically disturbed conditions shows that in temperate latitudes there is a clear depression of fE at noon as might be anticipated from considerations of the vertical drift associated with the disturbance current system. An analysis for an equatorial station indicates that at low latitudes magnetic disturbance may be accompanied by an enhancement of fE.

Certain other anomalies in fE in the neighbourhood of the auroral zone have also been briefly discussed.

Acknowledgements—We wish to thank the Astronomer Royal for supplying magnetic data for the Royal Greenwich Observatory, and Dr. YOSHIMATSU, the Director of the Kakioka Observatory, for similar data from certain Japanese observatories.

The above work was carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research, and this paper is published by permission of the Director of Radio Research.

REFERENCES

APPLETON E. V. 1955 *Proceedings of the Mixed Commission on the Ionosphere, Fourth Meeting*, p. 14.

APPLETON E. V. and LYON A. J. 1954 *Report of Conference on Physics of the Ionosphere*. p.20. Physical Society, London.

APPLETON E. V., LYON A. J. and TURNBULL A. G. 1955 *Nature, Lond.* **176,** 897.

APPLETON E. V., NAISMITH R. and INGRAM L. J. 1937 *Phil. Trans.* **236,** 191.

Appleton E. V., Naismith R. and Ingram L. J.	1939	*Proc. Phys. Soc. Lond.* **51,** 81.
Bartels J.	1949	*Terrestrial Magnetism and Electricity* (Edited by Fleming J. A.) p. 417. Dover Publications, New York.
Beynon W. J. G. and Brown G. M.	1956	*Nature, Lond.* **177,** 583.
Bibl K.	1951	*Ann. Géophys.* **7,** 208.
Cahill L. J. and Van Allen J. A.	1958	*J. Geophys. Res.* **63,** 270.
Chapman S.	1931	*Proc. Phys. Soc. Lond.* **43,** 483.
Chapman S.	1953	*Proc. Phys. Soc. Lond.* B **66,** 710.
Chapman S. and Bartels J.	1940	*Geomagnetism* p.229. Oxford University Press.
Endo H.	1957	*Mem. Kakioka Magn. Obs.* **8,** 1.
Findlay J. W.	1951	*J. Atmosph. Terr. Phys.* **1,** 353.
Greenhow J. S. and Neufeld E. L.	1956	*Phil. Mag.* Ser. 8, **1,** 1157.
Hasegawa M. and Ota M.	1937	*Proc. Imp. Acad. Japan* **13,** 69.
Huxley L. G. H.	1957	*Ann. I.G.Y.* **3,** Part III, 250.
Kirkpatrick C. B.	1948	*Aust. J. Sci. Res.* A **1,** 423.
McNish A. G.	1937	*IUGG Bull.* **10,** p. 271. Copenhagen.
Maeda K.	1955	*Rep. Ionos. Res. Japan* **9,** 60.
Martyn D. F.	1947a	*Proc. Roy. Soc.* **189,** 241.
Martyn D. F.	1947b	*Proc. Roy. Soc.* **190,** 273.
Martyn D. F.	1948	*Proc. Roy. Soc.* **194,** 429, 445.
Martyn D. F.	1952	*Proc. Conf. Ionosph. Physics* (Edited by Katz L. and Gerson N. C.) Part B, p. 69.
Martyn D. F.	1953	*Phil. Trans.* **246,** 306.
Sato T.	1955	*Rep. Ionos. Res. Japan* **9,** 205.
Sato T.	1957	*J. Geomagn. Geoelect., Kyoto* **9,** 57.
Scott J. C. W.	1952	*J. Geophys. Res.* **57,** 369.
Shimazaki T.	1957	*J. Radio Res. Labs., Japan* **4,** No. 15, 37.
Shimazaki T.	1958	*J. Radio Res. Labs., Japan* **5,** No. 19, 35.
Singer S. F., Maple E. and Bowen W. A.	1952	*Nature, Lond.* **170,** 1093.
Stagg J. M.	1928	*Geophys. Mem., Lond.* **5,** No. 42. Met. Office, London.
Vestine E. H., Laporte L., Lange I. and Scott W. E.	1947	*Publ. Carneg. Instn.* No. 580.
Wäldmeier M.	1944	*Helv. Phys. Acta* **17,** 168.

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 167 to 174. Pergamon Press Ltd. Printed in Northern Ireland

On the hypothesis that the mean geomagnetic field for parts of geological time has been that of a geocentric axial multipole

S. K. RUNCORN
Physics Department, King's College, Newcastle upon Tyne 1

(*Received* 26 *January* 1959)

Abstract—There are strong theoretical reasons for postulating that the mean geomagnetic field is symmetrical about the axis of the earth's rotation. The possibility that for parts of geological time the mean field has not been dipolar, but has had multipolar components, has not hitherto been discussed. Palaeomagnetic observations from North America, Australia and Europe are discussed in relation to this hypothesis but it is shown that continental drift is the more satisfactory inference to be drawn from the discrepancies between the polar wandering paths for the different continents.

VOL.
14
1959

1. INTRODUCTION

RUNCORN (1959) has given a simple argument in support of the basic proposition by which palaeomagnetic observations are being interpreted, i.e. that, when averaged over some thousands of years, the geomagnetic field is symmetrical about the axis of the earth's rotation. The further hypothesis that this axially symmetrical field has always been that of a geocentric dipole is more difficult to justify and in this paper we consider the consequences of postulating that the field has for considerable parts of geological time been that of an axial multipole.

CREER, IRVING and RUNCORN (1957) point out that the palaeomagnetic observations cannot be explained on this hypothesis alone for the declination of the mean field would always have been zero. However, it is worth examining whether a multipole field and a wandering pole of rotation are together sufficient to explain the differences between the polar wandering paths determined for Britain (CREER, IRVING and RUNCORN, 1954, 1957), North America (RUNCORN, 1956; COLLINSON and RUNCORN, 1959) and for Australia (IRVING and GREEN, 1958) without postulating continental drift.

2. MULTIPOLE FIELDS

RUNCORN (1954) argues that as the Coriolis force is dominant in the equation of fluid motion in the earth's core, a field permanently inclined to the axis of rotation is difficult to envisage. But RUNCORN (1959) shows that, even if it were possible for a permanent part of the geomagnetic field to be non-axial, the relative rotation of the core and the mantle would cause this to be averaged out when the field is observed on the earth's surface.

Now suppose that the mean geomagnetic field for appreciable parts of geological time has had multipole components. CREER, IRVING and RUNCORN (1957) point out, for example, that if the earth's mean field were a quadrupole the relation

between a palaeomagnetic angle of inclination (I) and the ancient colatitude (θ) of the site would be

$$\tan I = \cot \theta - \tfrac{1}{2} \tan \theta \tag{1}$$

instead of the familiar formula for a dipole field

$$\tan I = 2 \cot \theta \tag{2}$$

In general, the scalar potential V of a magnetic field of internal origin symmetrical about the earth's geographical axis can be expressed as the sum of terms

$$V = C_n a (a/r)^{n+1} P_n(\theta)$$

where C_n is the moment of the nth multipole, $P_n(\theta)$ the Legendre function of degree n and a is the radius of the earth. The vertical component (Z) of the earth's field (measured positive downwards) and the horizontal component (X) (measured positive northwards) are given by

$$Z = \frac{dV}{dr} = -C_n(n+1) P_n(\theta)$$

and

$$X = \frac{1}{r}\frac{dV}{d\theta} = C_n \frac{dP_n(\theta)}{d\theta}$$

It will be convenient for the purposes of this paper to depart from the usual sign convention for the angle of inclination (I) and take it to be measured from 0° to 360° with respect to the northward horizontal, counting clockwise, viewed from the east.

$$\operatorname{Tan} I = \frac{Z}{X} = -(n+1) P_n(\theta) \Big/ \frac{dP_n(\theta)}{d\theta} \tag{3}$$

thus I is zero along n lines of latitude.

Fig. 1 shows the variation of I between $\theta = 0$ and $\theta = 180°$ for the axial dipole, quadrupole and octupole fields.

3. Test of Multipole Hypothesis against Palaeomagnetic Data

To examine whether, in any geological period, the mean geomagnetic field had multipole components—assuming no continental drift—we proceed as follows. On the stereographic projection, meridians or great circles are drawn through the sites at which the palaeomagnetic observations for that period were made, and through the corresponding pole positions calculated on the dipole hypothesis. The discrepancies between these pole positions are now assumed to result from the incorrectness of the latter hypothesis and not from continental drift since the period in question. The true pole of the earth's rotation lies at an unknown distance θ from any site along its meridian. The new hypothesis breaks down at once if the meridians for one geological time are not concurrent, noting, that the position of the meridians may be varied, at the 95 per cent probability level, so long as they pass through the ovals of confidence. We now examine the data for

VOL. 14 1959

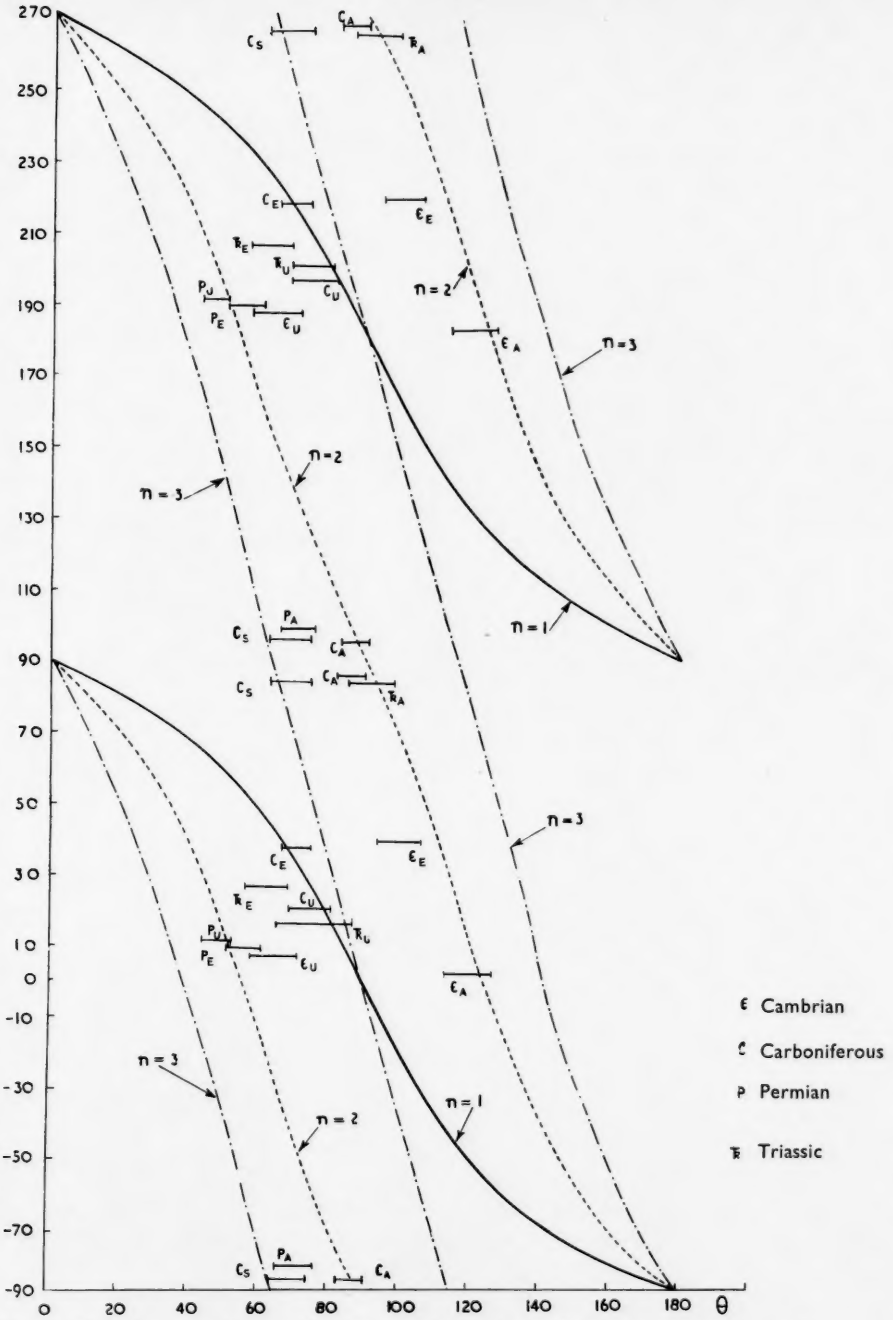


Fig. 1. Plot of angle of inclination (I) against colatitude (θ) for axial dipole ($n = 1$), quadrupole ($n = 2$), octupole ($n = 3$).
Suffixes: A denotes results from Australia. U from the United States and S from South Africa E from Europe.

the different geological periods obtained in Australia by Irving and Green (1959), in Europe by Creer, Irving and Runcorn (1957) and in North America by Runcorn (1956).

Figs. 2–6 show the sites and corresponding meridians for the late Pre-Cambrian, Cambrian, Carboniferous, Permian and Triassic. The late Pre-Cambrian meridians cannot be made to intersect at the 95 per cent probability level. However in this

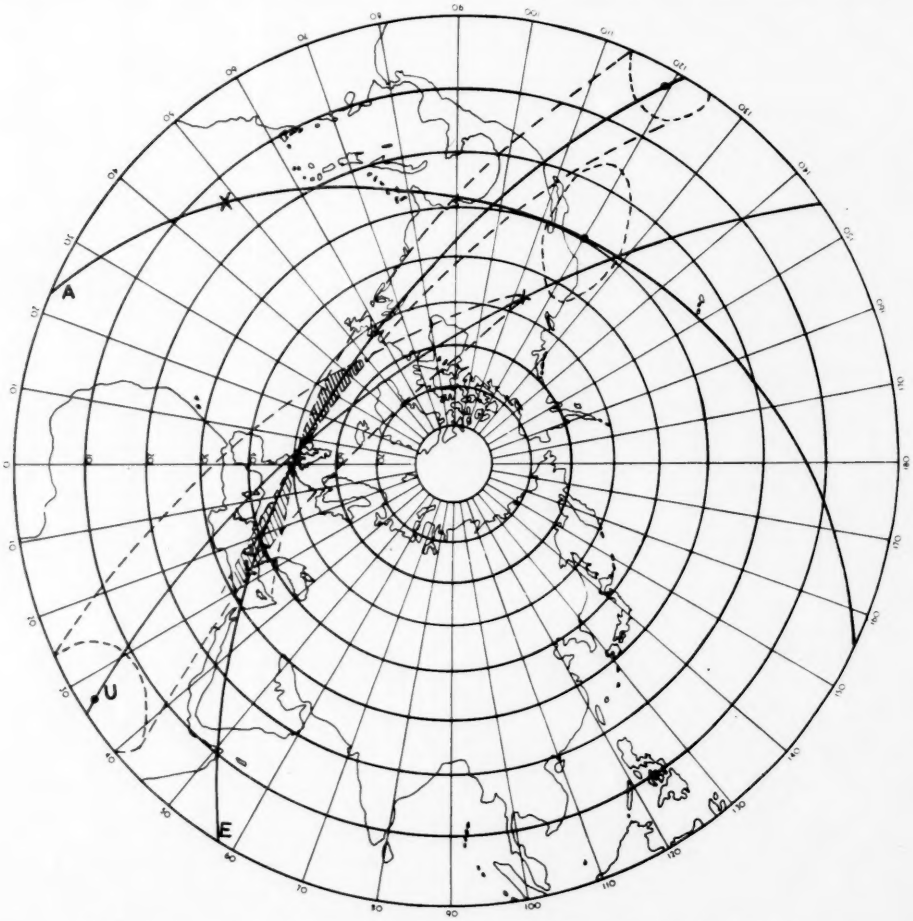


Fig. 2. Upper Pre-Cambrian meridians for Australia, Europe and North America.

part of the geological column there must be doubt as to the contemporaneity of the deposits. The meridians for the other times may be arranged to be concurrent, as shown, by varying their position within the allowed limits of error. The shaded areas are those in which the point of intersection, i.e. the supposed pole, may lie. Suppose now that the centre point of this area is regarded as the most probable position for the geographical axis on the multipole hypothesis, then the colatitudes θ of the sites in the various continents may be determined with respect to this pole.

These values are plotted on Fig. 1 against the value of I determined from the palaeomagnetic measurements: the horizontal lines are the spread corresponding to placing the pole at various positions within the shaded areas.

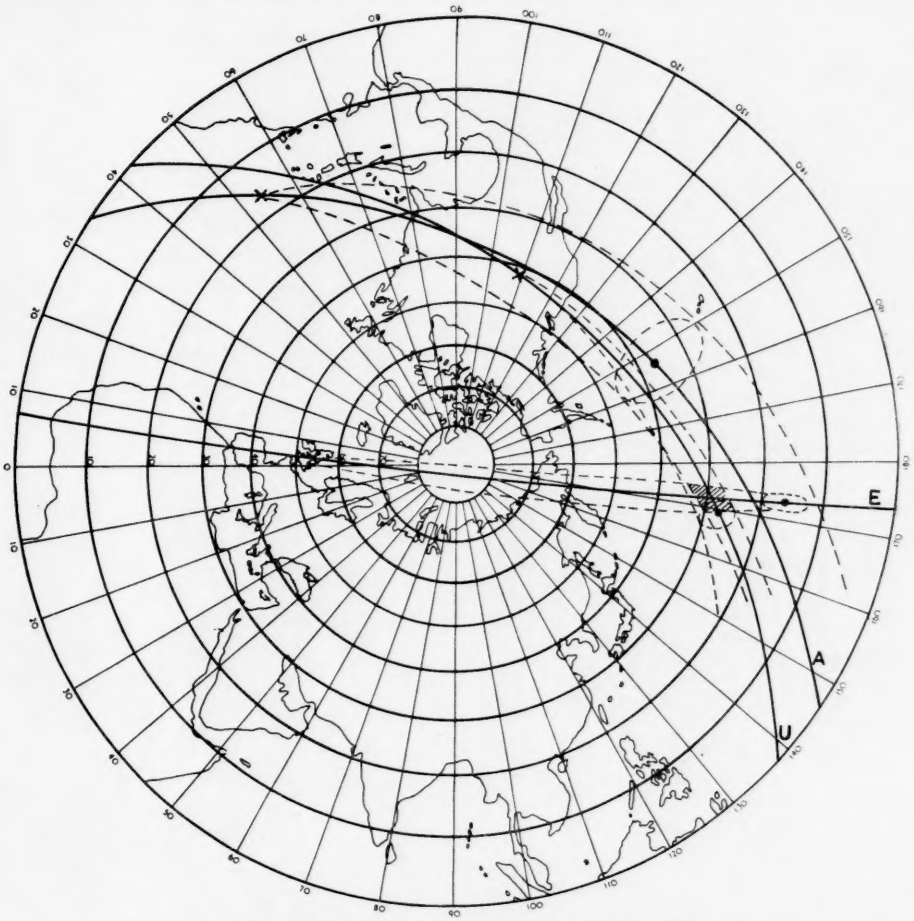


Fig. 3. Cambrian meridians for Australia, Europe and North America.

The following conclusions can be drawn:

(1) The Cambrian results for Australia, Europe and N. America are not well represented by an axial quadrupole field, very poorly represented by an axial octupole field and cannot be represented by an axial dipole field.

(2) The positions of the sites for the Carboniferous range from 30°N to 5°S of the equator, for the Permian from about 40°N to 15°N and for the Triassic from 35°N to 10°S with respect to the supposed poles. These positions are unreasonable on two grounds: (a) the angles of inclination for each period at these sites range from 0° to 90°. Consequently the mean geomagnetic field would have to be such that its angle of inclination changed more rapidly with latitude than any of the

VOL. 14 1959

curves shown. The Triassic and Carboniferous results come nearest to being satisfied by the octupole field and the Permian by the quadrupole field, but the discrepancy is still very distinct. (b) The positions of Australia in Carboniferous and Triassic times are within 10° of the equator, and the positions of South Africa

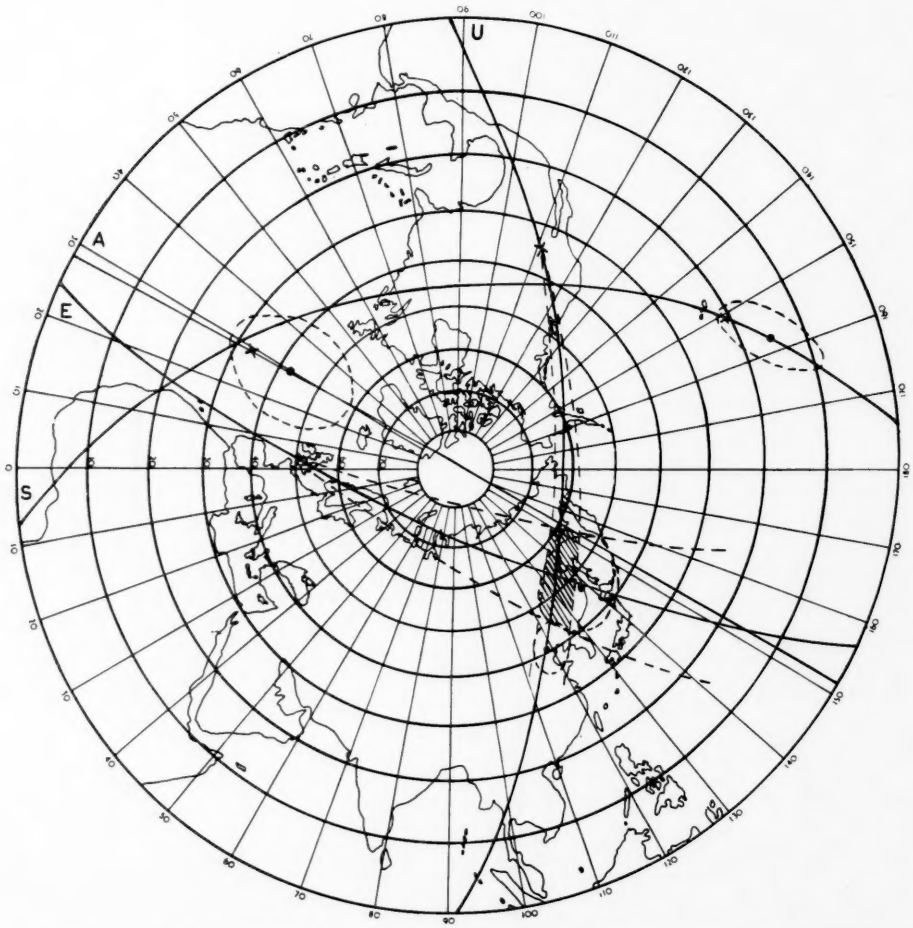


Fig. 4. Carboniferous meridians for Australia, Europe, North America and South Africa.

in Carboniferous times and of Australia in Permian times are within 30° of the equator. These positions are hard to reconcile with the evidence of a cool climate in Australia in Triassic times and widespread glaciations in Permo-Carboniferous times in Australia and South Africa (David, 1950; du Toit, 1939).

4. Conclusion

It is of course clear that the discrepancies between palaeomagnetic data from different continents, which have been taken by many writers to imply that

VOL
14
195

continental drift has occurred, can be explained away by postulating a sufficiently complex geomagnetic field, for each epoch in geological time. But the only physically plausible fields are axial multipoles and these seem to imply positions of the continents in flat contradiction to the palaeoclimatic data for late Palaeozoic

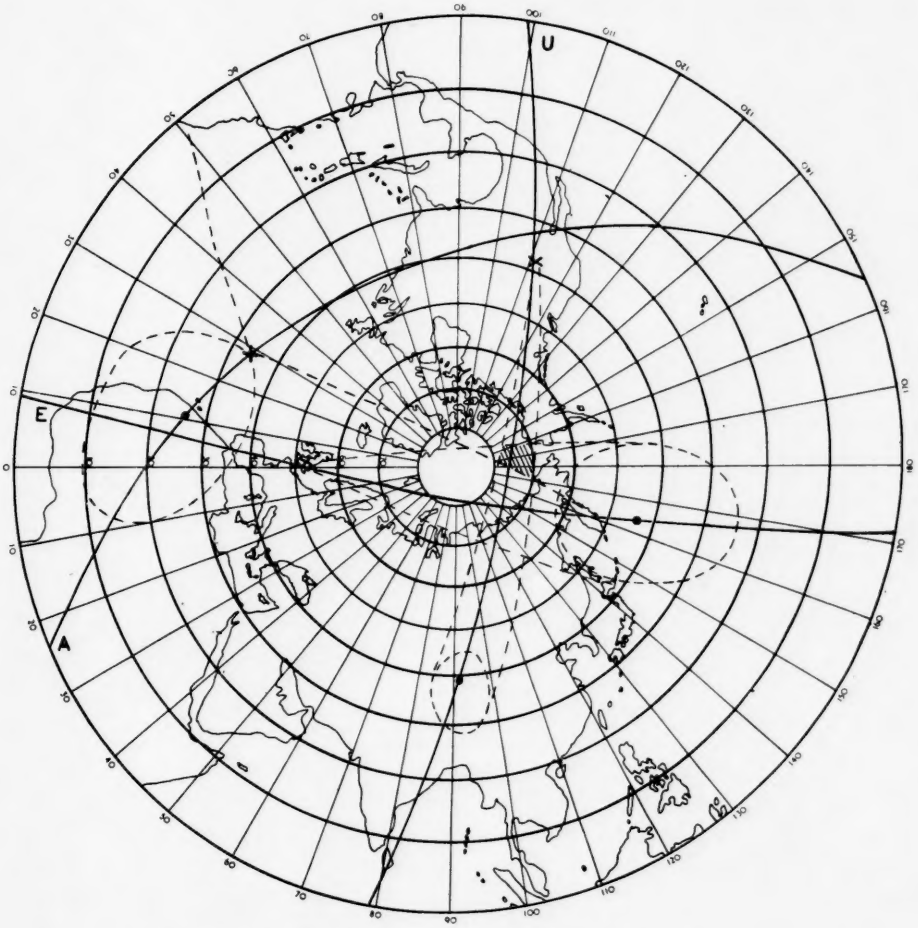


Fig. 5. Permian meridians for Australia, Europe and North America.

and early Mesozoic times. Until results from other continents are available the possibility cannot be entirely excluded that the mean geomagnetic field in Cambrian times was not dipolar.

Acknowledgements—The author is grateful to Mr. G. Turnbull for considerable help in the preparation of this paper, and to Mr. C. Helsley of the Geology Department, Princeton University, for a stimulating discussion of this hypothesis.

VOL. 14 1959

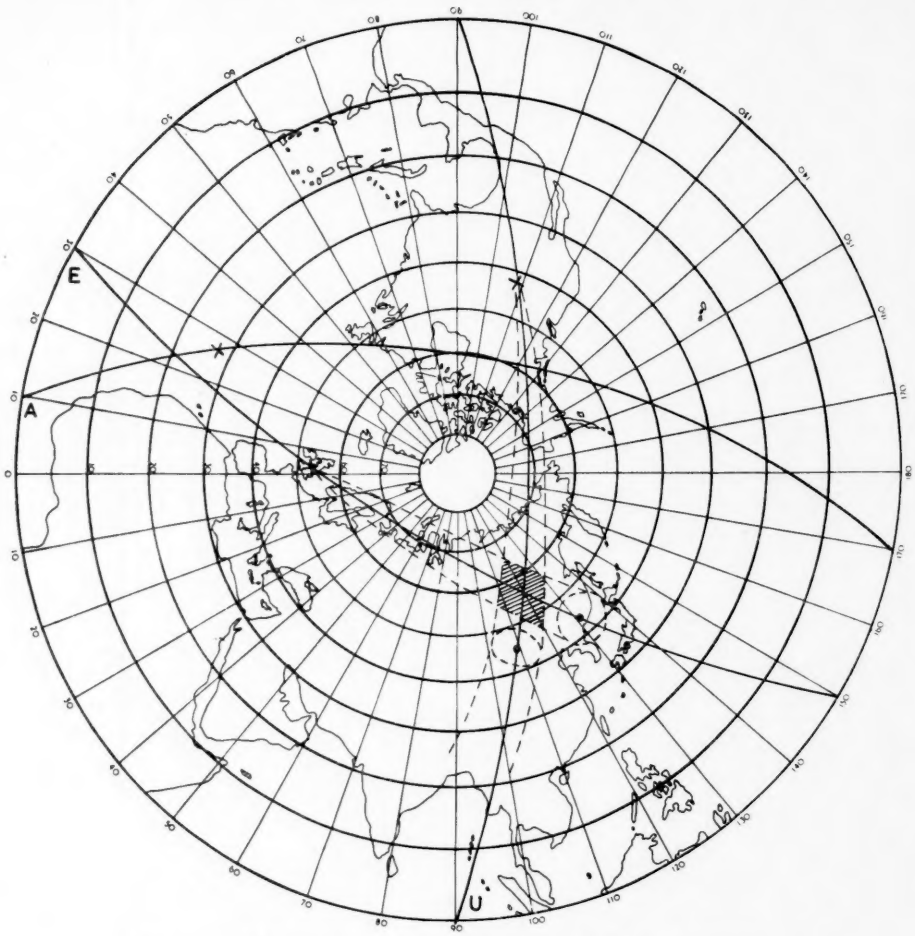


Fig. 6. Triassic meridians for Australia, Europe and North America.

References

Collinson D. W. and Runcorn S. K.	1959	*Bull. Geol. Soc. Amer.* In press.
Creer K. M., Irving E. and Runcorn S. K.	1954	*J. Geomagn. Geoelect., Kyoto* **6,** 163.
Creer K. M., Irving E. and Runcorn S. K.	1957	*Phil. Trans.* A **250,** 71.
David T. W. E.	1950	*The Geology of the Commonwealth of Australia.* Vol. I. Arnold, London.
du Toit A. L.	1939	*The Geology of South Africa.* 2nd Edition. Boyd & Carpenter, Edinburgh.
Nairn A. E. M.	1959	*Ov. Geol. Min. Res.* **7.** In press.
Runcorn S. K.	1954	*Trans. Amer. Geophys. Un.* **35,** 49.
Runcorn S. K.	1959	*Ann. Géophys.* **15,** 87.
	1956	*Bull. Geol. Soc. Amer.* **7,** 301.

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 175 to 180. Pergamon Press Ltd. Printed in Northern Ireland

RESEARCH NOTES

A rapid method of obtaining accurate virtual heights from an ionogram

(*Received* 26 *August* 1958)

Abstract—A transparent slider is described which corrects the virtual height on an ionogram for changes due to varying echo amplitude and which enables accurate virtual heights to be obtained rapidly.

THE standard methods of measuring the virtual height of a layer from the trace on an ionogram are inherently inaccurate owing to the variation of the apparent virtual height with the amplitude of the reflected signal. The height observed shows both systematic and variable errors which are usually about ± 5 km but can reach 15 km in unfavourable circumstances. As has been shown by LYON and MOORAT (1956), these errors can be corrected provided that the trace on the ionogram shows characteristics which vary

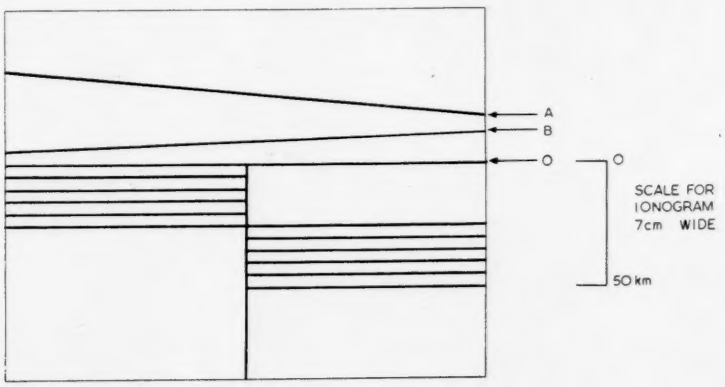


Fig. 1. Layout of slider for reducing virtual height accurately.

regularly with the echo amplitude. This condition is usually satisfied automatically when the output signal has not been differentiated. The width of the dead zone following a differentiated pulse may often be used as the amplitude-sensitive parameter when a differentiator is in use but the correction law often proves to be more complicated and less reliable than for undifferentiated signals.

For most ionosondes the relation between the width of the trace, ω, and the error in vertical height, Δ, is linear:

$$\Delta = a - b\omega \tag{1}$$

the constants a, b depending on the design of the instrument and the pulse width used. For panoramic-type equipments, a and b are constants but most slow-speed type recorders

show a small increase in a at the lower frequencies. Some typical values for the DSIR Mark II Recorders are $b = 0{\cdot}25$, $a = 14{\cdot}5$ for frequencies above 2 Mc/s and $b = 0{\cdot}25$, $a = 17{\cdot}5$ for frequencies below 2 Mc/s, Δ, ω being measured in kilometres of virtual height.

In order to reduce ionograms rapidly, it is desirable to be able to identify the correct virtual height corresponding to a trace of any width. This may be done conveniently by using a transparent overlay which gives, in effect, a graphical solution of equation (1). A suitable design is shown in Fig. 1. In practice the frequency at which the height is being measured can be most easily identified by adding a secondary cursor, with a vertical line, which can slide along the overlay but this is not essential.

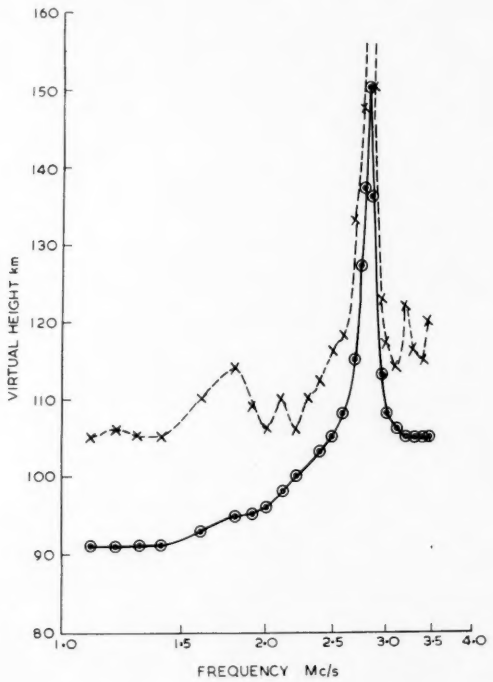


Fig. 2. Ionogram taken at Slough 0900 GMT, 16 July 1954. × Virtual height from actual trace. ⊙ Virtual height using slider.

The constants for the equipment are determined using the methods described by Lyon and Moorat (1956) and the two sloping lines A, B, are drawn so that when the vertical distance between A and B represents ω km the vertical distance between A and the horizontal line O represents Δ km. The two sets of auxiliary lines below O subdivide the interval between the height calibration lines on the ionogram and thus simplify accurate measurement.

In practice the overlay is moved until the upper and lower edge of the trace at the required frequency just touch the sloping lines A, B. The line O then indicates the corrected position of the lower edge of the trace and the relative position of this line and the height calibration gives the required virtual height. After a little practice the time taken does not differ significantly from that needed to measure the uncorrected height although accuracies of about 1 to 2 km are now readily available for ionograms of suitable size.

As with all height measurements, difficulties can arise when overlapping or scattered echoes are present. In general, only the lowest of several overlapping traces can be measured accurately and it may not always be possible to identify the trace width corresponding to the echo concerned. It is usually best to assume that the pulse width does not vary greatly with frequency when the zone of confusion is limited in frequency. In the absence of any help from the record, one must adopt the width corresponding to the most probable amplitude. This is usually close to the average width found when the echo patterns are simple. In many cases, however, inspection of the ionogram shows the most probable width for the main trace, scattered echoes or noise causing local displacements of the front or rear edge of the trace which clearly do not obey equation (1). It must not be forgotten that the heights deduced when scattered echoes are predominant may differ appreciably from those obtained when the ionosonde is operated at a lower gain level which eliminates the weaker traces immediately below and above the main trace. It is therefore more important to monitor the behaviour of the recorder when accurate height measurements are required. A comparison of ionograms made at high and low gain levels has shown that when the upper edge of the trace is varying fairly smoothly with frequency but the lower edge is confused by weaker reflections, suitably corrected measurements of the upper edge of the trace, assuming a normal trace width, give results which are remarkably consistent with those found from the low gain record. This situation is common when measuring E-region traces.

A typical virtual height variation with frequency, obtained quickly by this method, using an ionogram with a height scale of 100 km in 2 cm, is shown in Fig. 2. The uncorrected heights obtained directly from the record are also shown. The differences in the height and shape of the curves indicate the importance of using a correct method of reduction.

Acknowledgement—The work described above was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

W. R. Piggott

Official Communication from
DSIR Radio Research Station, Slough, Bucks

VOL. 14 1959

Reference

Lyon A. J. and Moorat A. J. G. 1956 *J. Atmosph. Terr. Phys.* **8**, 309

Cross-sections for K-shell ionization of atomic nitrogen and oxygen by electrons and protons

(*Received* 1 *November* 1958)

1. Introduction

In another connexion a general study has recently been made of K-shell ionization of atoms by electrons and protons. The results for nitrogen and oxygen are presented here since they are of interest in aeronomy.

2. Electron Impact

The cross-section for K-shell ionization by electrons of energy E, velocity v and momentum $\hbar\mathbf{k}_1^0$ is given in the relativistic Born approximation of Møller (1932) by

$$Q = (\gamma/Z_e\mathbf{k}_1^0)^2 S\pi a_0^2$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$, Z_e is the effective nuclear charge and S is a quantity which, for a given value of $\mathscr{E} = E/E_0$ above 2·5, is almost independent of Z_e and of ϑ, the ratio of the actual K-shell ionization potential E_0 to the fictitious ionization energy Z_e^2 Rydbergs which would arise in the absence of outer screening (cf. Arthurs and Moiseiwitsch 1958).

The published S values are for an $\mathscr{E}$-range up to 20. This range has been extended to

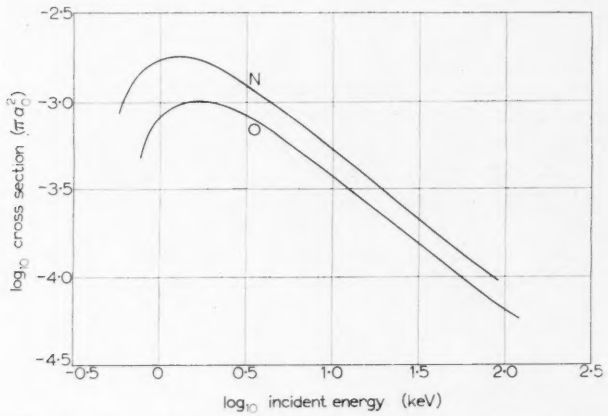


Fig. 1. Cross-sections for K-shell ionization of N and O by electrons $[\pi a_0^2 = 0{\cdot}88 \times 10^{-16}\ \mathrm{cm}^2]$.

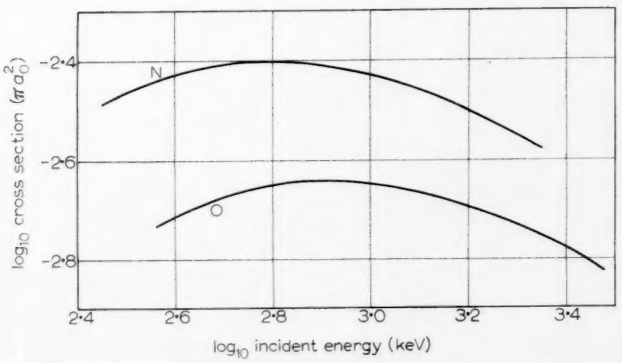


Fig. 2. Cross-sections for K-shell ionization of N and O by protons $[\pi a_0^2 = 0{\cdot}88 \times 10^{-16}\ \mathrm{cm}^2]$.

225 by using the Born cross-sections for ionization of atomic hydrogen cited by McCarroll (1957) for $\mathscr{E}$ up to 72·25 and by computing the remaining values from the Bethe (1930) approximation, which fits on to the Born approximation in this region. Table 1 contains the S values from which relativistic cross-sections may readily be derived. The cross-sections for nitrogen and oxygen are presented in Fig. 1.

VOL.
14
1959

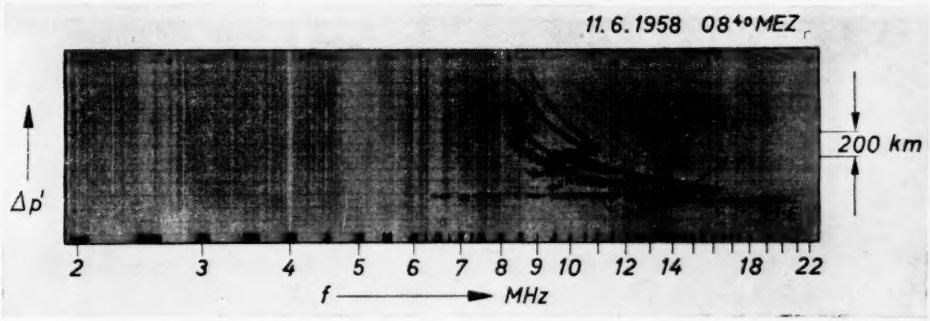


Fig. 1. Oblique-incidence ionogram over a distance of 2000 km (Sodankylä–Lindau).

VOL
14
195

Table 1

$\mathscr{E}$	25	36	49	72·25	100	121	144	169	196	225
S	16·8	17·8	18·6	19·8	20·4	20·9	21·3	21·6	22·0	22·3

The cross-section for K-shell ionization is given by $Q = \left[\left(\frac{\gamma^2}{\gamma^2 - 1}\right) \frac{S}{(137Z_e)^2}\right] \pi a_0^2$ and $\mathscr{E} = E/E_0$, where E is the impact energy of the incident electron and E_0 is the experimental ionization energy of the relevant atom.

3. Proton Impact

Calculations with the Born formula have been performed on a Deuce computer for proton impact, taking ϑ equal to 0·65 which is close to the values appropriate for nitrogen ($\vartheta = 0{\cdot}65_5$) and oxygen ($\vartheta = 0{\cdot}66_3$). The Stobbe–Massey–Mohr limits of integration were used in preference to those of Bethe (cf. Arthurs, 1958). Fig. 2 contains the results.

Acknowledgements—The writer wishes to thank Professor D. R. Bates, F.R.S., for his guidance and interest in these calculations.

The research reported has been sponsored in part by the Air Force Cambridge Research Center, Geophysics Research Directorate, of the Air Research and Development Command, U.S. Air Force, through its European Office, under contract No. AE61(052)-131.

Department of Applied Mathematics,
The Queen's University of Belfast

A. M. Arthurs

VOL. 14 1959

References

Arthurs A. M.	1958	Submitted for publication.
Arthurs A. M. and Moiseiwitsch B. L.	1958	*Proc. Roy. Soc.* A **247,** 550.
Bethe H.	1930	*Ann. Phys., Lpz.* **5,** 325.
McCarroll R.	1957	*Proc. Phys. Soc. Lond.* A **70,** 460.
Møller C.	1932	*Ann. Phys., Lpz.* **14,** 531.

Further results of sweep-frequency oblique-incidence pulse transmissions

It is common practice in calculating MUF's to consider the *F2*- and *E*-reflections only, and to disregard the effect of *F1* because for many circuits, and for undisturbed conditions, the *F1*-MUF is less than the *F2*- or *E*-MUF. By sweep-frequency oblique-incidence pulse experiments it can be shown that over a path of 2000 km in northern Europe the omission of the *F1*-reflections leads to an erroneous concept of the propagation conditions. As a matter of fact, in summertime during most of the daylight period the junction point ("nose") between the low- and high-angle ray of the *F1*-path is observed at higher frequencies than that of the *F2*-path (Fig. 1). On undisturbed days at noon the difference is 1–2 Mc/s in favour of the *F1*-MUF increasing to 5 Mc/s on slightly disturbed days. The daily

variation during a day with a minor geomagnetic disturbance is shown in Fig. 2. The *F1*-MUF overcomes the *F2*-MUF at 0430 hours and falls behind at 1800 hours.

It is interesting to note that the irregular variations of the *F2*-MUF (a) are accompanied by similar variations of the virtual height of the nose of the *F2*-path. (b) Calculating the

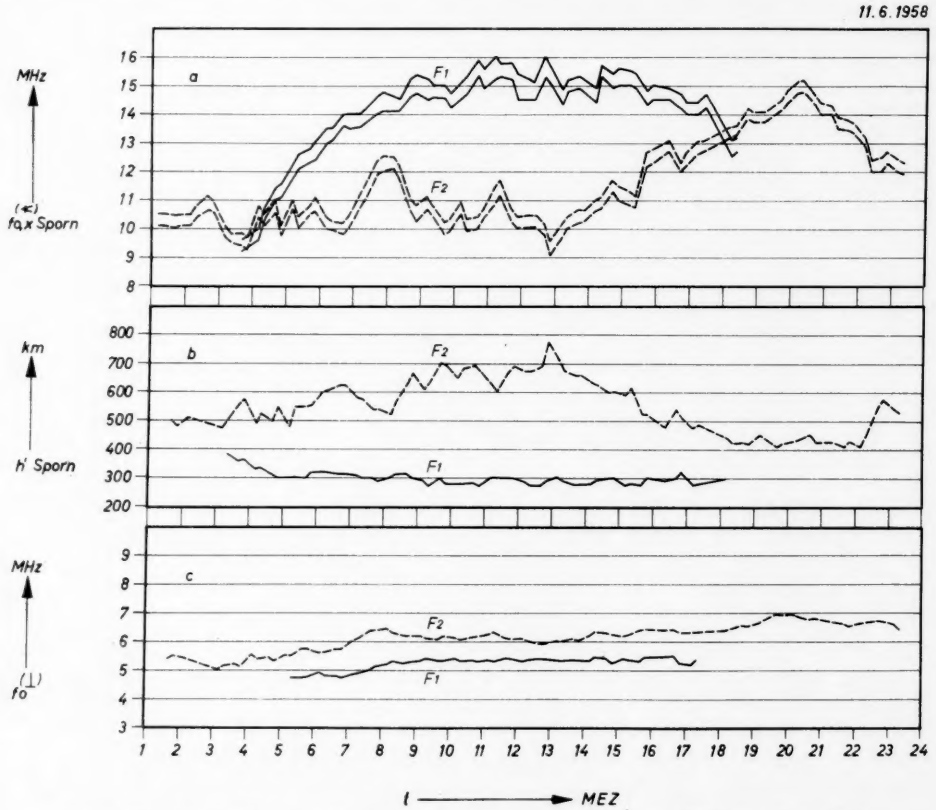


Fig. 2. Results of oblique-incidence observations over a distance of 2000 km (Sodankylä–Lindau). (a) Observed *F2*- and *F1*-MUF (*x* and 0); (b) observed virtual height of *F1*- and *F2*-nose; (c) calculated equivalent critical frequency of *F2* and *F1*.

equivalent critical frequencies at vertical incidence (c) it becomes evident that simultaneously the latter varies rather smoothly with time. So we may conclude that the strong variations of the *F2*-MUF are caused by variations of the virtual height (due to variable retardation in *F1* probably) rather than by changes of the critical frequency.

A detailed analysis will be given in a later communication.

W. Dieminger
H. G. Möller
G. Rose

Max-Planck-Institut für Aeronomie,
Institut für Ionosphären-Physik,
(206) *Lindau über Northeim (Hann.), Germany*

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 181 to 182. Pergamon Press Ltd. Printed in Northern Ireland

Indices of geomagnetic activity

for the Observatories HARTLAND (*Ha*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

October and November 1958

The figures given on pages 181 to 182 represent the *K*-indices for three-hour intervals, beginning with 00–03 hrs for the first and ending with 21–24 hrs for the eighth figure.

October 1958

Day	*Ha* Range for $K = 9 : 500\gamma$		*Es* Range for $K = 9 : 750\gamma$		*Le* Range for $K = 9 : 1000\gamma$		*Day*
	K-Indices	*Sum*	*K-Indices*	*Sum*	*K-Indices*	*Sum*	
1	1343 2234	22	1333 3235	23	1233 2235	21	1
2	3332 1113	17	3322 2113	17	3422 2113	18	2
3	2433 3442	25	2332 3331	20	2332 3552	25	3
4	1121 2002	9	1011 2212	10	1111 2212	11	4
5	4122 3223	19	3123 3333	21	4222 2333	21	5
6	1212 3233	17	1213 3233	18	1222 3223	17	6
7	2222 2334	20	2322 3334	22	2222 2324	19	7
8	3322 1113	16	3322 2112	16	3322 1112	15	8
9	1111 2111	9	1112 2211	11	1112 1211	10	9
10	0011 1211	7	1111 2221	11	1111 1211	9	10
11	2121 2123	14	2122 1123	14	2111 1222	12	11
12	1011 1111	7	0001 1211	6	1111 1211	9	12
13	0123 1221	12	0113 2221	12	1112 2211	11	13
14	1212 3212	14	1212 3212	14	1212 3112	13	14
15	3222 2222	17	2122 2222	15	2122 2122	14	15
16	1322 2111	13	2322 2211	15	2222 1211	13	16
17	1221 2222	14	1222 2332	17	2222 1223	16	17
18	2321 2212	15	2221 2212	14	2221 2212	14	18
19	1222 2322	16	1112 2322	14	1212 2322	15	19
20	3111 1123	13	2011 2223	13	3111 1213	13	20
21	1221 1111	10	2211 1111	10	1222 2111	12	21
22	1443 3555	30	1443 3555	30	1443 3676	34	22
23	4443 3454	31	4442 4444	30	5455 4566	40	23
24	3454 4655	36	4453 4655	36	5654 6677	46	24
25	3232 1211	15	3232 1211	15	6232 2211	19	25
26	0122 1242	14	1121 2243	16	1122 2342	17	26
27	3211 1365	22	3212 2364	23	3322 2586	31	27
28	3244 4444	29	3244 4444	29	5244 4555	34	28
29	3222 2354	23	3322 1354	23	5232 2255	26	29
30	3223 2335	23	3223 3334	23	4122 3344	23	30
31	2222 2322	17	2222 2412	17	2212 1321	14	31

November 1958

Day	*Ha* *Range for* $K = 9 : 500\gamma$			*Es* *Range for* $K = 9 : 750\gamma$			*Le* *Range for* $K = 9 : 1000\gamma$			*Day*
	K-Indices		*Sum*	*K-Indices*		*Sum*	*K-Indices*		*Sum*	
1	1022	3222	14	1132	3222	16	1122	3112	13	1
2	1012	2542	17	1112	2543	19	1012	3652	20	2
3	3322	3122	18	3322	3222	19	3322	2222	18	3
4	2213	4132	18	2213	3133	18	2313	3133	19	4
5	0010	0000	1	1011	1111	7	1111	1111	8	5
6	0000	0002	2	0011	2112	8	1111	1112	9	6
7	0012	2212	10	1112	2212	12	1012	2313	13	7
8	0111	1101	6	0112	1111	8	1112	1111	9	8
9	0221	0101	7	0222	2211	12	0122	2110	9	9
10	0223	2332	17	1223	2332	18	0123	2332	16	10
11	4421	2110	15	4332	2111	17	5321	2111	16	11
12	0102	3232	13	1212	3232	16	1102	2132	12	12
13	2211	1101	9	2221	1111	11	2311	1111	11	13
14	1000	0103	5	2110	2103	10	1000	1112	6	14
15	2121	2001	9	2111	2100	8	2111	2111	10	15
16	2222	1201	12	2322	2201	14	1222	2201	12	16
17	1222	1212	13	1222	2212	14	1122	1112	11	17
18	1021	2233	14	1021	2233	14	1121	2244	17	18
19	1101	1100	5	1111	1200	7	1111	2201	9	19
20	1101	1100	5	1111	2100	7	1111	1100	6	20
21	1001	0101	4	1011	2200	7	1011	1200	6	21
22	1101	1000	4	1100	1000	3	1101	1100	5	22
23	1102	3201	10	1002	2111	8	2111	2101	9	23
24	0012	3222	12	0012	3222	12	0022	2222	12	24
25	0110	2333	13	1111	2333	15	0101	2434	15	25
26	3201	2232	15	3111	1232	14	3212	2332	18	26
27	1211	2332	15	1212	2332	16	1212	2332	16	27
28	3333	3222	21	3333	3222	21	3332	3112	18	28
29	3312	2301	15	3222	1311	15	3322	2311	17	29
30	0001	1000	2	0101	1000	3	0101	0000	2	30

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 183 to 184. Pergamon Press Ltd. Printed in Northern Ireland

Papers to be published in future issues

G. McK. Allcock: The electron density distribution in the outer ionosphere derived from whistler data

R. Montalbetti: Photoelectric measurements of hydrogen emissions in aurorae and airglow

C. M. Minnis and G. H. Bazzard: Some indices of solar activity based on ionospheric and radio-noise measurements

C. H. Cummack: Chapman behavior in the lower ionosphere

G. S. Kent: High-frequency fading observed on the 40 Mc/s wave radiated from artificial satellite 1957α

G. S. Aitchison and K. Weekes: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—I. The theory of the analysis

G. S. Aitchison *et al.*: Some deductions of ionospheric information from the observations of emissions from satellite 1957α 2—II. Experimental procedure and results

C. A. Ventrice and E. R. Schmerling: Coefficients for the rapid reductions of h'–F records to N–h profiles without computing aids

E. L. Hagg *et al.*: Spiral occurrence of sporadic-E

F. Hepburn: Interpretation of smooth-type atmospheric wave forms

M. W. McElhinny: Some further analyses of E-layer measurements in South Africa during the solar eclipse of 25 December 1954

C. S. Warwick and R. T. Hansen: Geomagnetic activity following large solar flares

W. H. Ward: On the origin of terrestrial particles from solar flares

V. A. Bailey: Some possible effects caused by strong gyro waves in the ionosphere—I

M. H. Rees: Absolute photometry of the aurora—I

M. H. Rees: Absolute photometry of the aurora—II

P. C. Clemmow and M. A. Johnson: A contribution to the theory of the motion of weak irregularities in the ionosphere

G. F. Walton: Calculation of the vertical distribution of atmospheric ozone

N. F. Barber and D. D. Crombie: VLF reflections from the ionosphere in the presence of a transverse magnetic field

L. G. H. Huxley: A discussion of the motion in nitrogen of free electrons with small energies, with reference to the ionosphere

J. M. Malville: Type-B aurorae in the Antarctic

W. Becker: New methods and some results concerning true ionospheric height calculation

P. Bandyopadhyay: The early morning $E2$-layer and some evidence of pre-sunrise F-layer splitting

W. L. Price: Anomalous ionospheric reflection during solar eclipses

J. D. Whitehead: Sudden changes in the virtual height of radio waves reflected from the E-region of the ionosphere

M. D. Vickers: The effect of the $F1$-layer on the calculation of the height of the $F2$-layer

F. Mariana: A new theoretical model of the composite F-layer

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—I. Solar variations

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—II. Lunar and luni-solar variations in H and Z

E. Woyk: The refraction of radio waves by a spherical ionized layer

A. J. Ferraro and J. J. Gibbons: Polarization by means of the multislab approximation

B. N. Singh and R. L. Ram: Rhythmic fading of short-wave radio signals

A. G. McNamara and D. W. R. McKinley: The effect of trail irregularities on the interpretation of meteor echoes

K. C. Clark and A. E. Belon: Spectroscopic observations of the great aurora of 10 February 1958—I. Abnormal vibration of N_2^+

K. C. Clark and A. E. Belon: Spectroscopic observations of the great aurora of 10 February 1958—II. Unusual atomic features

M. L. Meeks and J. C. James: Meteor-radiant distributions and the radio-echo rates observed by forward scatter

E. A. Lytle and D. M. Hunten: The ratio of sodium to potassium in the upper atmosphere

J. F. Noxon, A. W. Harrison and A. Vallance Jones: The infrared spectrum of the night airglow from 1.4 μ to 4.0 μ

G. Kvifte: Auroral and nightglow observations at Ås, Norway

H. K. Paetzold: The fluctuations of the acceleration of satellites

R. L. Ingraham: Note on the tidal theory of the S_q magnetic field

J. D. Whitehead: The absorption of short radio waves in the *D*-, *E*- and *F*-regions of the ionosphere

J. A. Fejer and R. W. Vice: An investigation of the ionospheric *D*-region

J. A. Fejer and R. W. Vice: The use of full-wave solutions in the interpretation of ionospheric absorption measurements

J. Egedal: On a pseudo-iterating method for reduction of observations of earth-tide for effects from tides of the sea, with an application to observations made at Bergen (Norway) 1934

K. Bibl, A. Paul and K. Rawer: Die frequenzabhängigkeit der ionosphärischen absorption

D. M. Schlapp: Frequency in the *E*-region of the ionosphere

G. V. Groves: A theory for determining upper-atmosphere winds from radio observations of meteor trails

J. M. Kelso: The determination of the electron density in interplanetary space

J. A. Gledhill: The effects of a solar eclipse on a stratified ionosphere

J. A. Gledhill: The behaviour of the ionosphere over Cape Town and Johannesburg during the annual solar eclipse

S. R. Khastgir: Study of 'winds' in the *F*-region of the ionosphere during the unusual days in the IGY calendar

J. S. Greenhow and E. L. Neufeld: Turbulence at altitudes of 80–100 Km and its effects on long-duration meteor echoes

H. J. A. Chivers and J. S. Greenhow: Auroral ionization and the absorption and scintillation of radio stars

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 185 to 199. Pergamon Press Ltd. Printed in Northern Ireland

The electron density distribution in the outer ionosphere derived from whistler data

G. McK. Allcock
Dominion Physical Laboratory, Department of Scientific and Industrial Research, Lower Hutt, New Zealand

(*Received* 16 *July* 1958)

Abstract—A method, using successive approximations, is given for finding the variation of electron density N with height h in the outer ionosphere, when the variation of whistler dispersion with geomagnetic latitude is already known. From the application of the method to specific whistler data, it is inferred that the electron density distribution in the outer ionosphere decreases exponentially with height at least in the height range 1000 to 13,000 km, and that during the period January to May 1957 the equation describing the distribution within this height range was:

$$N \fallingdotseq 5{\cdot}75 \times 10^4 \exp(-h/2640)$$

where N is in cm^{-3} and h in km.

VOL. 14 1959

1. Introduction

It is now widely recognized that the analysis of whistlers can yield information on the properties and behaviour of the outer ionosphere to heights of several earth's radii. In particular, electron densities can be determined from whistler characteristics, and some approaches to this problem have already been made, e.g. Storey (1953, 1957) and Helliwell (1956).

The analytical determination of the distribution of electron density N with height h above the surface of the earth is beset by comparable difficulties, whether one is interested in deriving an (N, h) distribution for the normal ionospheric layers from ionograms, or for the outer ionosphere from whistler spectrograms, as in both cases the experimental data is given in the form of frequency/time diagrams. Because of these difficulties, empirical methods are used in the analysis of ionograms, with emphasis lying in the search for the method giving the best approximation to the true (N, h) distribution with the least computational effort. Thus it would be expected that the analysis of whistler spectrograms would proceed in the same way.

In this paper one possible empirical method of determining electron densities is presented. The method uses data on the variation of whistler dispersion with latitude. Since the latitude determines the maximum height of the path followed by a whistler, it is possible to quickly calculate a maximum value of electron density at any height. By using an iterative process this approximation is improved. The (N, h) distribution has been derived by this method from whistler data for the period January to May 1957. It is thought that this represents to within 50 per cent the average electron density distribution between heights of 1000 and 1300 km during this period.

2. Method of Analysis

2.1. *Initial assumptions*

Various initial assumptions have been made, as follows. The validity of these assumptions is discussed later. Throughout this paper the term "latitude" always refers to geomagnetic latitude.

(a) The ray path for all frequency components of a whistler is assumed to lie exactly along a line of force of the earth's magnetic field, which in turn is assumed to be a dipole field.

(b) The experimental data is assumed to be given as the variation of dispersion D with latitude l_0 of the lightning-flash source. Here D is defined according to the Eckersley–Storey approximation (Storey, 1953):

$$D \fallingdotseq t_f f^{1/2} \tag{1}$$

where t_f is the time of travel of the whistler component of frequency f during a single transit from one hemisphere to the other, i.e. for a short whistler.

It has already been assumed in the derivation of equation (1) that f is much less than either the gyrofrequency f_H or the critical (or plasma) frequency f_0 everywhere along the whistler path. Under these conditions Storey has shown that

$$D \fallingdotseq \frac{1}{2c} \int \frac{f_0}{f_H^{1/2}}\, ds \tag{2}$$

where c is in the velocity of light *in vacuo*, and ds is an element of path length, the integration being taken over the whole path. By using the relationships $f_0^2 = Ne^2/\varepsilon_0 \pi m$ and $f_H = \mu_0 eH/2\pi m$ (where e, m are the electronic charge and mass, and ε_0, μ_0 the electric and magnetic permittivities of free space respectively, H being the magnetic field strength) equation (2) leads to the expression

$$D \fallingdotseq \left(\frac{e}{2}\right)^{1/2} \int \left(\frac{N}{H}\right)^{1/2} ds \tag{3}$$

where e is in e.m.u., H in G, N in cm^{-3} and ds in cm, thus giving D in the usual unit of sec$^{1/2}$.

(c) It is further assumed that surfaces of constant electron density are spherical about the centre of the earth, and that the electron density decreases monotonically with height throughout the region under consideration. The lower limit to this region is taken as the level of maximum ionization in the *F2*-layer. The upper limit is determined by the height above the geomagnetic equator of the magnetic line of force through the highest source latitude. This height is found from the general equation:

$$h = a \tan^2 l_0 \tag{4}$$

where a is the radius of the earth.

2.2. *Upper limit to electron density*

On the above assumptions, an upper limit N_1 can be quickly set to the value of electron density at any height. If the electron density were *constant* with height, then from equation (3)

$$D \fallingdotseq N^{1/2} \left(\frac{e}{2}\right)^{1/2} \int \frac{1}{H^{1/2}}\, ds \tag{5}$$

Now, for a uniformly magnetized earth or its equivalent centred dipole, the

magnetic field strength at a point of latitude l, distant r from the centre of the earth, is given by

$$H = \frac{M(1 + 3 \sin^2 l)^{1/2}}{r^3} \tag{6}$$

where M is the magnetic moment of the earth. But the (r, l) equation for a line of force is

$$r = a \frac{\cos^2 l}{\cos^2 l_0} \tag{7}$$

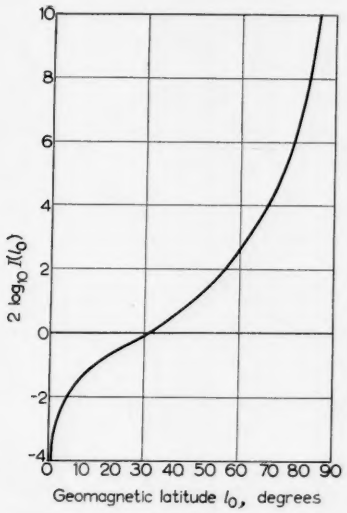


Fig. 1. Curve showing the effect on whistler dispersion of the earth's magnetic field as a function of geomagnetic latitude, used when determining the upper limit of electron density from equation (10).

Combining equations (5), (6) and (7) eventually leads to:

$$D \fallingdotseq N^{1/2} \left(\frac{e}{2}\right)^{1/2} \left(\frac{2a^{5/2}}{M^{1/2}}\right) \left[\frac{1}{\cos^5 l_0} \int_0^{l_0} \cos^4 l (4 - 3 \cos^2 l)^{1/4} \, dl\right] \tag{8}$$

Denoting the quantity in square brackets by $I(l_0)$, and inserting numerical values for e, a and M, equation (8) becomes

$$D \fallingdotseq 0{\cdot}202 \, N^{1/2} I(l_0) \tag{9}$$

Because of the great variation of $I(l_0)$ with l_0, it is useful to take (common) logarithms and rearrange equation (9) as

$$\log N \fallingdotseq 1{\cdot}389 + 2 \log D - 2 \log I(l_0) \tag{10}$$

The quantity $2 \log I(l_0)$ is shown as a function of l_0 in Fig. 1. From this graph, used in conjunction with equation (10), the value of constant electron density which would produce the measured dispersion can be rapidly calculated. Now, on the assumption made initially that in fact N decreases monotonically with h, it is clear that the value of N calculated from equation (10) must be greater than the actual value at the top of the path. The value of N thus calculated can therefore be

VOL. 14 1959

regarded as an upper limit N_1 to the electron density at the height h given by equation (4). The curve drawn through the values of (N_1, h) will be taken as a first approximation to the true (N, h) distribution, bearing in mind that N may actually be much less than N_1.

2.3. *Method of successive approximations*

From this point the method consists in deriving successively closer approximations. A second approximation is obtained by taking the (N_1, h) curve found above, and proceeding to calculate a set of new dispersion values (D_1, l_0). For a

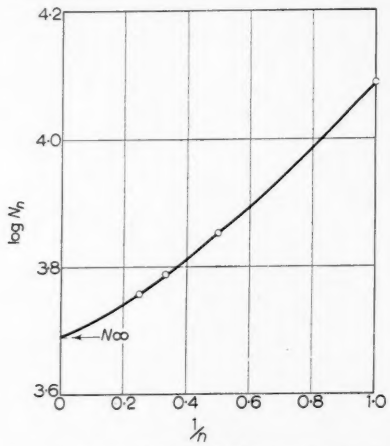


Fig. 2. Curve showing the graphical method of estimating the final approximation ($N\infty$) from early approximations.

given l_0, D_1 will be necessarily greater than the experimental value D, if the assumption of monotonically decreasing electron density with height is correct. Since D is roughly proportional to $N^{1/2}$, a better value N_2 for the electron density at a given height will be

$$N_2 = \left(\frac{D}{D_1}\right)^2 N_1 \tag{11}$$

In this way an (N_2, h) curve is calculated which is a closer approximation to the actual distribution than that given by the (N_1, h) curve. But, since the ratio (N_2/N_1) from equation (11) will increase with h and thus with l_0, the correction given by equation (11) will be insufficient, and the (N_2, h) curve will lie in between the (N_1, h) and the true (N, h) curve. Hence successive approximations by this method will approach the true (N, h) curve always from the same direction.

Clearly the $(n + 1)$th approximation is obtained from the nth approximation by the generalized form of equation (11), viz.

$$N_{n+1} = \left(\frac{D}{D_n}\right)^2 N_n \tag{11a}$$

In practice, after three or four approximations, it is possible to accelerate the process by plotting N_n or, more profitably $\log N_n$ against $(1/n)$ and extrapolating by eye to infinite n, as in Fig. 2, thus obtaining what, for convenience, will be

called N_∞. A check calculation of D_∞ as a function of l_0 can then be compared with the experimental (D, l_0) data. The differences $(D_\infty \sim D)$ should be less than the original experimental error; if not, one further approximation using equation (11a) will almost certainly suffice.

2.4. *Method of calculating* (D, l_0) *curves*

In the above method, most of the computing time is taken in calculating the (D, l_0) curve from the assumed (N, h) curve. A fairly quick approximate method of doing this was developed, based on the use of CHAPMAN and SUGIURA's tables of magnetic dipole field arc lengths (CHAPMAN and SUGIURA, 1956).

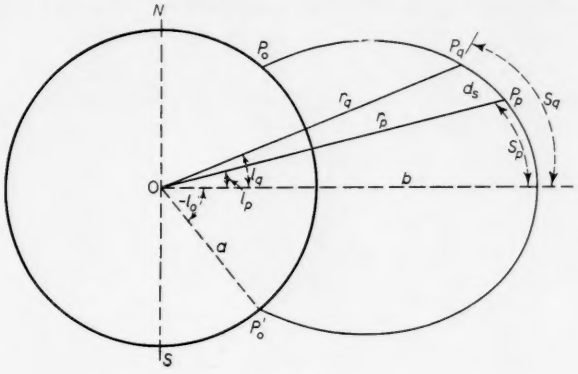


Fig. 3.

In Fig. 3, consider the elemental contribution to the dispersion within the interval P_pP_q along the magnetic line of force which enters or leaves the earth at P_0 (or P_0') whose latitude is l_0 (or $-l_0$). Let the interval P_pP_q be small enough so that the quantity $(N/H)^{1/2}$ varies almost linearly throughout this interval. Then the elemental contribution to the dispersion is given by the differential form of equation (3), viz.

$$dD \fallingdotseq \left(\frac{e}{2}\right)^{1/2} \left(\frac{N}{H}\right)^{1/2} ds \tag{12}$$

where the average value of $(N/H)^{1/2}$ over the interval P_pP_q is

$$\frac{1}{2}\left\{\left(\frac{N_p}{H_p}\right)^{1/2} + \left(\frac{N_q}{H_q}\right)^{1/2}\right\}$$

and $ds = s_q - s_p = b\{f(l_q - f\,(l_p)\}$ where b is a function of l_0 only and $f(l)$ is a function of l only:

$$b = a + h = a \sec^2 l_0$$

$$f(l) = \frac{1}{2\sqrt{3}}\,(x + \sinh x \cosh x)$$

$$\sinh x = \sqrt{3} \sin l$$

b and $f(l)$ are tabulated by CHAPMAN and SUGIURA.

VOL. 14 1959

Thus equation (12) can be re-written as

$$dD \doteq \left(\frac{b}{2}\right)\left(\frac{e}{2}\right)^{1/2}\left\{\left(\frac{N_p}{H_p}\right)^{1/2} + \left(\frac{N_q}{H_q}\right)^{1/2}\right\}\{f(l_q) - f(l_p)\} \tag{13}$$

Corresponding to the dispersion contribution from the interval P_pP_q, there is an equal contribution from the interval $P_p'P_q'$, the mirror image of P_pP_q about the equatorial plane. Therefore the total dispersion over the path P_0QP_0' is found by summing contributions of the form of equation (13) over the latitude range 0 to l_0 and doubling the answer, i.e.

$$D \doteq b\left(\frac{e}{2}\right)^{1/2}\sum_{0}^{l_0}\left\{\left(\frac{N_p}{H_q}\right)^{1/2} + \left(\frac{N_q}{H_p}\right)^{1/2}\right\}\{f(l_q) - f(l_p)\} \tag{14}$$

In practice it has been found that little error is introduced if the intervals P_pP_q are derived from equal increments of r, each of 1000 km for source latitudes higher than 40°, and of 500 km for lower source latitudes. In making a numerical calculation of D from equation (14) for a given latitude, many of the steps in the calculation are universally applicable to that latitude, irrespective of the electron density distribution. Thus, once D has been calculated for a given (N, h) distribution, additional calculations required for other (N, h) distributions are limited merely to the few involving N.

3. Results

The method outlined above was applied to experimental whistler data obtained at the following stations during the period January to May 1957:

Wellington, N.Z. (geomagnetic latitude—45·1°)
Washington, D.C. (50·2°)
Dunedin, N.Z. (—51·0°)
Hanover, N.H. (55·1°)

In addition, data from two Japanese stations, Toyokawa (24·5°) and Wakkanai (35·3°) published by Iwai and Outsu (1956), were used to provide lower latitude information. Unfortunately the Japanese data does not cover an identical period, the Toyokawa data referring to February 1956 and the Wakkanai data to July–August 1956. From the relationship between whistler dispersion and sunspot number found by Allcock and Morgan (1958), it is estimated that this may result in the (N, h) curve being depressed by 30–40 per cent at heights of 1000–3000 km.

A further source of error is the way in which the data from the American and New Zealand stations were obtained. Here dispersions have been inferred from aural determinations of the time delay between the atmospheric bonk arriving directly from the lightning source, and the first frequency heard in the resulting whistler, according to equation (1). This frequency was normally about 5 kc/s (where maximum whistler energy is received) and is believed to have been close to this value, on the average. Occasional check spectrograms have confirmed this, to within 10 per cent.

The experimental dispersion data are plotted in Fig. 4 against latitude,

VOL
14
195

together with an additional abscissa giving the maximum height associated with each latitude from equation (4). This assumes that the average whistler was produced by a lightning source at the same latitude as that of the station.

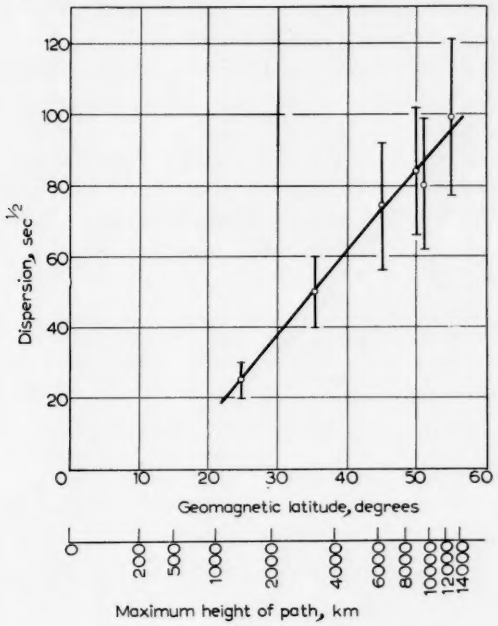


Fig. 4. Experimental values of dispersion plotted against geomagnetic latitude and height of path above equator.

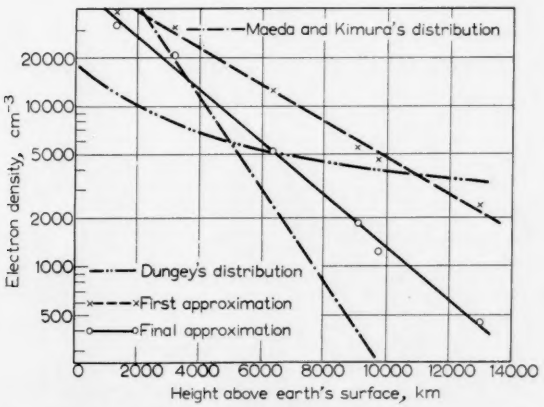


Fig. 5. The (N, h) curve derived from the data of Fig. 4 by the method of this paper, together with other (N, h) distributions for comparison.

The first ("upper limit") approximation to the true (N, h) distribution was obtained by applying equation (10) to the data from each station in turn. This approximation is shown in Fig. 5. It is clear that the plotted points marked "X"

are adequately represented by a straight line when log N and h are taken as the co-ordinates. The process of successive approximations was then carried out according to the method outlined above, with one minor modification: the contribution to the total dispersion arising in the regular ionospheric layers was assumed to be equivalent to that obtained by allowing the (N, h) distribution under consideration to extend down to a height of 100 km, and assuming that the ionization below this level was negligible. The uncertainty arising from this procedure is not likely to be more than 2 or 3 per cent.

The calculations were taken to the fourth approximation before estimating the final approximation in the manner already described. The final approximation is shown as the full curve in Fig. 5. That this closely represents the true (N, h) distribution appears likely from a comparison of the experimental dispersion data with the dispersions calculated from the final approximation, in Table 1.

Table 1. Comparison of experimental values of whistler dispersion with those calculated from the (N, h) distribution of Fig. 5 (full curve)

Station	Experimental dispersion ($sec^{1/2}$)	Dispersion calculated from final (N, h) approximation ($sec^{1/2}$)
Toyokawa	25 (± 5)	26
Wakkanai	50 (± 10)	46
Wellington	74 (± 18)	75
Washington	84 (± 18)	85
Dunedin	80 (± 18)	88
Hanover	99 (± 22)	97

From this it is inferred that the electron density distribution between heights of 1000 and 13,000 km for the period January to May 1957 is represented adequately by the equation

$$N \fallingdotseq 5{\cdot}75 \times 10^4 \exp\left(\frac{-h}{2640}\right) \qquad (15)$$

where N is in cm^{-3} and h in km. Thus, within this height range, the outer ionosphere appears to have an exponentially decreasing electron density with a scale height of 2640 km. For this data at present available, it is not necessary to postulate a more complex model of the outer ionosphere than this.

4. Discussion of Assumptions

This section discusses briefly some of the assumptions made in Section 2.

4.1. *Ray path*

Maeda and Kimura (1956) have shown that, for the particular electron density distribution which they have taken, a considerable deviation of the ray path outwards from a magnetic line of force can take place, so that the bulk of the whistler energy arising from a lightning source may arrive in the opposite hemisphere at a

point up to 2000 km nearer the geomagnetic pole than the conjugate point. The extent of this outward deviation (or "polar creep") is, however, critically dependent upon the electron density distribution, especially that near the top of the path. The distribution of equation (15) shows a much greater electron density at extreme heights than does that taken by MAEDA and KIMURA, which is also shown in Fig. 5. The amount of polar creep is therefore expected to be much less than that calculated in their paper. Other experimental evidence to elucidate this point is somewhat meagre. The rare occurrence of trains of whistler echoes in which the amplitudes of successive echoes increases, reaches a peak, and later decreases, has been interpreted as being due to polar creep. On the other hand DINGER (1956) notes that at Washington whistlers are heard from storms in the north more frequently than from storms in the south, in spite of the fact that storms are generally more numerous in the south. This would certainly not be expected if polar creep were important.

4.2. *Shape of earth's magnetic field*

Recent determinations of the "cosmic ray equator" (e.g. ROSE *et al.*, 1956) have shown that it is displaced from the geomagnetic equator by several degrees in places. Thus the shape of the earth's magnetic field evidently differs from that of a simple dipole field in the outer ionosphere as well as near the surface of the earth. From the analysis of surface measurement of the field, a higher-order approximation to the true field has been obtained by BLOCK and HERLOFSON (1956) by postulating a quadrupole in addition to the dipole. They have calculated the positions of the southern hemisphere points conjugate to 123 selected points in the northern hemisphere, taking into account the quadrupole term. Some support for these calculations can be obtained by comparing the occurrences of whistlers at Wellington and Dunedin and at Unalaska, Alaska. On the dipole approximation, the point conjugate to Unalaska would be distant 750 km from Wellington, but only 250 km from Dunedin. The inclusion of the quadrupole term shifts the conjugate point to a position only 280 km from Wellington, but 430 km from Dunedin, i.e. it is now nearer to Wellington than to Dunedin. For the period 3 November 1956 to 3 February 1957, it was found that on 46 per cent of the occasions when a short whistler was heard at Unalaska, a corresponding long whistler was heard at Wellington, whereas on only 20 per cent of such occasions was a corresponding long whistler heard at Dunedin.

It appears from BLOCK and HERLOFSON's calculations that the lines of force through Hanover and Washington re-enter the earth about 5° nearer the south geomagnetic pole, and those through Dunedin and Wellington 2–3° nearer the north geomagnetic pole, than would the simple dipole field; so that in all four cases the whistler path is somewhat longer than was assumed in the calculations of Section 3. For this reason the derived values of electron density in equation (15) may be 20 to 30 per cent higher than actual.

4.3. *Distribution of electron density with latitude*

It was initially assumed that surfaces of constant electron density were spherical about the centre of the earth, i.e. that the (N, h) distribution is independent of

latitude. This is obviously an idealization, since the total electron content per unit column up to the *F2*-layer peak is known to be greater at the equator than at the poles. On the other hand, there is presumably a greater probability of electrons from extra-terrestrial sources entering the outer ionosphere at high latitudes. In the absence of quantitative data this question of spherical symmetry remains unanswered. It may be that a detailed examination of whistler spectrograms from a range of known source latitudes, in comparison with curves of the type shown in Fig. 7, discussed in the next section, will reveal the extent to which the (N, h) distribution does in fact vary with latitude.

5. The Nose Whistler Effect

A more exact treatment of whistler propagation leads to a further correction of the electron densities in equation (15). It has been pointed out by Helliwell *et al.* (1956) and independently by Ellis (1956) that for high latitude whistlers, the gyrofrequency $f_{H\,\min}$ at the top of the whistler path is not much greater than the highest received whistler frequency. For instance if the source latitude is 60°, then $f_{H\,\min}$ is 13·8 kc/s. Under these circumstances, the Eckersley–Storey approximation no longer holds, and instead of finding a time of travel t_f from equations (1) and (2), viz.

$$t_f \doteqdot \frac{1}{2c}\int \frac{f_0}{f^{1/2}f_H^{1/2}}\,ds \tag{16}$$

a modified time t_f' is obtained from

$$t_f' \doteqdot \frac{1}{2c}\int \frac{f_H f_0}{f^{1/2}(f_H - f)^{3/2}}\,ds \tag{17}$$

in which f is now allowed to approach f_H in magnitude, so long as it is still much less than f_0 (e.g. if $N = 400$ cm^{-3}, then $f_0 = 178$ kc/s).

Expressing equations (16) and (17) in differential form, and combining them, we have

$$dt_f' = \frac{1}{[1 - (f/f_H)]^{3/2}}\,dt_f \tag{18}$$

This shows that the time spent in travelling through a small length ds of the path is greater than would be expected on the Eckersley–Storey approximation, by a factor $[1 - (f/f_H)]^{-3/2}$ which will here be called the "nose whistler function" $F(f/f_H)$. This function is shown in Fig. 6; it approaches unity as f approaches zero, so that when f is much less than f_H the Eckersley–Storey approximation holds. The dispersion D as already defined in equation (1) according to this approximation may thus be identified with a fictitious "dispersion at zero frequency" D_0, i.e.

$$D_0 = \operatorname*{Lt}_{f\to 0}\,(t_f f^{1/2}) \tag{19}$$

At any other frequency, if equation (1) is used, a measurement of the travel time t_f' will yield an "apparent dispersion" at that frequency (D_f'), which will be greater than D_0 since, from equation (18), dt_f' is greater than dt_f throughout the ionospheric path.

It can be seen from the above that t_f' will become infinitely large as f approaches

either f_H or zero. For some value f_N of frequency lying between 0 and f_H, the value of t_f' will be a minimum; f_N is called the "nose frequency" and is the value of satisfying the condition

$$\frac{dt_f'}{df} = \frac{1}{2c}\frac{d}{df}\left[\int \frac{f_H f_0}{f^{1/2}(f_H - f)^{3/2}}\, ds\right] = 0 \tag{20}$$

Spectrograms of high-latitude whistlers showing this nose effect have been given by HELLIWELL *et al.* and by DINGER (1956). Our concern here is, generally to calculate the spectrograms to be expected from the (N, h) distribution derived in the

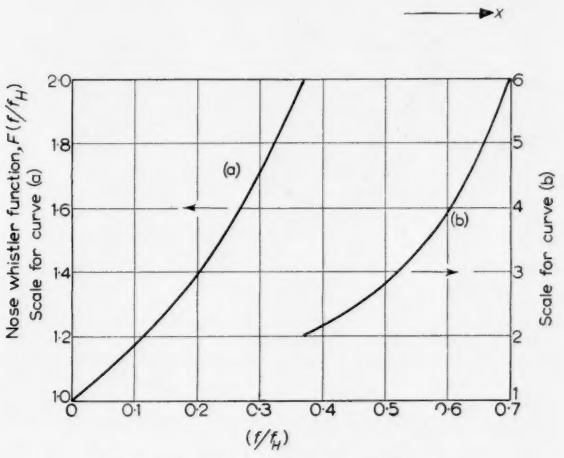


Fig. 6. The nose whistler function
$F(f/f_H) = (1 - f/f_H)^{-3/2}$

previous section, and particularly, to determine to what extent the derivation, which was based on the Eckersley–Storey approximation, was valid at 5 kc/s.

The apparent dispersion D_f', and consequently the travel time t_f', is calculated by mutiplying the right-hand side of equation (13) by the nose whistler function averaged over the interval $P_p P_q$, i.e.

$$dD_f' \fallingdotseq \left(\frac{b}{4}\right)\left(\frac{e}{2}\right)^{1/2}\left\{F_p\left(\frac{f}{f_H}\right) + F_q\left(\frac{f}{f_H}\right)\right\}\left\{\left(\frac{N_p}{H_p}\right)^{1/2} + \left(\frac{N_q}{H_q}\right)^{1/2}\right\}\{f(l_q) - f(l_p)\} \tag{21}$$

and summing as before. As this equation is dependent upon frequency, the summation must be made separately for a number of frequencies in order to plot the spectrogram. However, since the Eckersley–Storey approximation calculations using equation (14) will have already been made, little extra computing effort is needed to determine the extent of the nose effect.

The calculated spectrograms for source latitudes of 45°, 50° and 55° are shown as full curves in Fig. 7. These latitudes are closely those of Wellington, Washington and Hanover respectively. The broken curves represent the corresponding Eckersley–Storey approximations. A few spectrograms have been made of

VOL. 14 1959

whistlers received at Wellington; these are consistent with the curves of Fig. 7 up to the highest analysed frequency (8·5 kc/s), when it is taken into account that no information on the location of the sources was available, and therefore the source latitude could have differed from the receiver latitude by 10° or more.

In conclusion, the curves of Fig. 7 show that the nose effect at 5 kc/s introduces

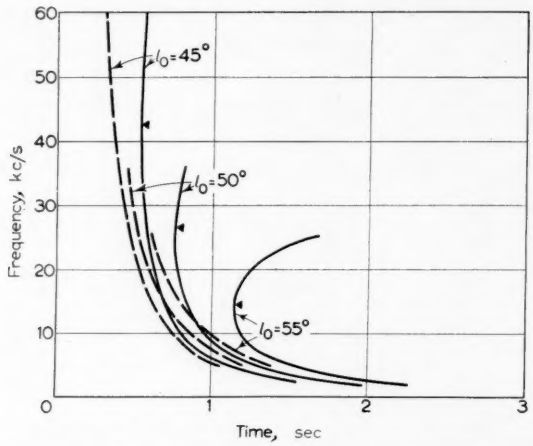


Fig. 7. Spectrograms for latitudes 45°, 50° and 55° calculated from the (N, h) distribution of Fig. 5. (The triangles denote the nose frequencies; the dotted curves are the corresponding spectrograms on the Eckersley–Storey approximation.)

about 10 per cent additional travel time for $l_0 = 55°$, and lesser amounts at lower latitudes. Hence the (N, h) distribution of equation (15) may indicate electron densities up to 20 per cent higher than actual at extreme heights (in the vicinity of 13,000 km) because of this.

6. The Electron Density at Heights Above 13,000 km

The frequent occurrence has been noted of multiple whistlers apparently arising from the same source. At high latitudes, where f_N is within the recorded range of frequencies, multiple nose whistlers have been observed, in which f_N decreases for successive components. This phenomenon can be interpreted as follows:

A lightning flash illuminates the lower edge of the ionosphere over a range of latitudes up to 10° or more on either side of the flash. Under certain conditions, probably associated with magnetic disturbances, the lower edge of the ionosphere can be considered as the location of several secondary sources, each launching a pulse of energy into the ionosphere at its appropriate latitude. In this way series of whistlers are produced at the conjugate region, each whistler having been guided along a separate path, and acquiring values of f_N and t_{fN}' characteristic of the latitude of its secondary source. Since f_N decreases and t_{fN}' increases with latitude, in the resulting series of whistlers, f_N is found to decrease with travel-time.

If we allow this interpretation, we can then proceed to calculate from Fig. 7 how f_N would vary with travel-time. By using equation (15) this curve can be extrapolated to higher latitudes. To facilitate comparison with experimental data, the

rate of decreases of f_N is plotted against f_N in Fig. 8. The appropriate values of l_0 are marked on the curve, and the region covered by the extrapolation is shown as a broken curve. For comparison, such experimental data as have been published by HELLIWELL and by DINGER are also plotted. There is close agreement for the points in the vicinity of $l_0 = 55°$, but for higher latitudes the experimental rate of decrease of f_N fairly quickly falls below the extrapolated rate. It can be shown

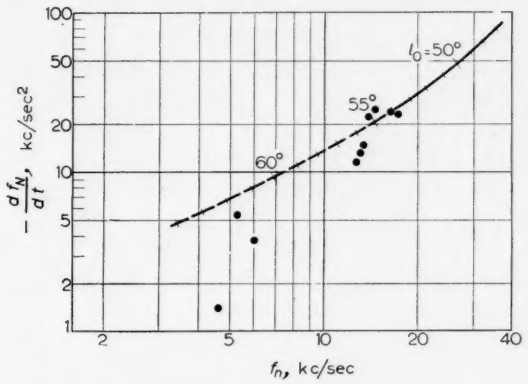


Fig. 8. Curve derived and extrapolated from Fig. 7 showing the rate of decrease of nose frequency in a multiple whistler as a function of the nose frequency, compared with experimental points.

VOL. 14 1959

that this would happen if the electron density decreased less rapidly with height above 13,000 km than is given by equation (15). Such information as is available for greater heights (e.g. HELLIWELL, 1956) suggests that the electron density at 27,000 km may in fact be from ten to twenty times that calculated by extrapolating equation (15) beyond the range of the data on which it is based.

7. THE DUNGEY DISTRIBUTION OF IONIZATION

DUNGEY (1954) has suggested on theoretical grounds that the electron density distribution in the outer ionosphere may be of the form

$$N = N_0 \exp\left(\frac{2{\cdot}5}{R}\right)$$

where R is the distance from the centre of the earth, measured in earth's radii. This distribution was used by STOREY (1957) as the starting-point for a worked example illustrating his method of determining the (N, h) distribution from a single whistler spectrogram. His discussion of likely errors arising from the use of his method is based on those arising in the worked example, in the hope that they will be "fairly representative of those likely to occur in practice". It is appropriate here to compare the (N, h) distribution derived in this paper with the Dungey distribution, plotted in Fig. 5. The value of N_0 in the Dungey distribution has been chosen so that the values of N are the same at a height of one earth's radius. The difference between the two distributions is immediately apparent.

A consequence of this difference is that the constant-density approximation

(i.e. the "upper limit" of Section 2.2, which is identical with Storey's "zeroth approximation") yields an electron density at a height of 13,000 km which is about five times that given by the final approximation, whereas for the Dungey distribution Storey finds that the constant-density approximation comes within 30 per cent of the true density at the top. The contribution to the dispersion is actually much more uniform throughout the height range than is given by the Dungey distribution.

8. Conclusion

An empirical method has been given for determining the variation of electron density with height in the outer ionosphere, when data on whistler dispersions are available from a number of stations, so that the variation of dispersion with latitude is initially known. The method has been applied to dispersion data from two North American and two New Zealand stations, averaged over the period January to May 1957, and supplemented by some earlier low-latitude data from two Japanese stations. Between the heights of 1000 and 13,000 km the derived (N, h) distribution is closely that of an ionosphere whose electron density decreases exponentially with height according to the equation

$$N \fallingdotseq 5{\cdot}75 \times 10^4 \exp\left(\frac{-h}{2640}\right)$$

where N is in cm^{-3} and h in km.

In the method various assumptions and approximations have been used, and the validity of these has been discussed. From this and from a few supplementary experimental results involving the nose whistler effect, it is estimated that the above equation is not likely to be in error by more than 50 per cent throughout this region for the period under consideration.

Acknowledgements—This work forms part of the radio research programme of the Dominion Physical Laboratory. The data for Hanover and Washington were kindly supplied by Dr. M. G. Morgan of Dartmouth College, Hanover, whilst the Dunedin station was operated by Mr. L. H. Martin of New Zealand Broadcasting Service. The author also acknowledges the benefit of helpful discussions with colleagues.

References

Allcock G. McK and Morgan M. G.	1958	*J. Geophys. Res.* In press.
Block L. and Herlofson N.	1956	*Tellus* **8,** 210.
Chapman S. and Sugiura M.	1956	*J. Geophys. Res.* **61,** 485.
Dinger H. E.	1956	*Report* 4825. Naval Research Laboratory, Washington, D.C.
Dungey J. W.	1954	*The Physics of the Ionosphere* p. 229. Physical Society, London.
Ellis G. R.	1956	*J. Atmosph. Terr. Phys.* **8,** 338.
Helliwell R. A.	1956	*Low Frequency Propagation Studies* Part 1. Radio Propagation Laboratory, Stanford, Calif.
Helliwell R. A., Crary J. H., Pope J. H. and Smith R. L.	1956	*J. Geophys. Res.* **61,** 139.
Iwai A. and Outsu J.	1956	*Proc. Res. Inst. Atmosph., Nagoya Univ.* **4,** 29.

VOL
14
195

Maeda K. and Kimura I.	1956	*Rep. Ionos Res. Japan* **10,** 105.
Rose D. C., Fenton K. B., Katzman J. and Simpson J. A.	1956	*Canad. J. Phys.* **34,** 968.
Storey L. R. O.	1953	*Phil. Trans.* A **246,** 113.
Storey L. R. O.	1957	*Canad. J. Phys.* **35,** 1107.

VOL. 14 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 200 to 212. Pergamon Press Ltd. Printed in Nothern Ireland

Photoelectric measurements of hydrogen emissions in aurorae and airglow*

R. MONTALBETTI†
University of Saskatchewan, Saskatoon, Canada

(*Received* 18 *July* 1958)

Abstract—Observations are reported on hydrogen emissions in aurorae and airglow made with a scanning spectrometer during the period 10–24 February 1958 at Churchill, Manitoba. No systematic relationship between the occurrence of the H_β line and the usual auroral forms was noted, with the exception of very strong active displays in which hydrogen emissions were not observed. Measurements of the intensity and profile of the H_β line indicate that a proton flux of some 10^7–10^8 protons/cm^2 sec, with some protons having an initial velocity of $\geq$3000 km/sec, are associated with some auroral displays. Strong hydrogen emissions, along with enhanced airglow radiations, were detected during an interval of no visible aurorae. Assuming that light scattering played an important role in the measurements, the H_α and H_β line profiles can be explained by an influx of monoenergetic protons having an initial velocity of about 1200 km/sec. The flux for the beginning of the enhanced radiations was calculated to be the order of 10^8 protons/cm^2 sec.

1. INTRODUCTION

SINCE the discovery of hydrogen Balmer emissions in some aurorae by VEGARD (1939), a number of researchers have continued to make measurements in an attempt to assess the importance of the impinging proton stream as a primary excitation source for the aurora. The Doppler shift and the line broadening reported by MEINEL (1951) and VEGARD (1956) leave little doubt that, on occasions, the protons do enter the upper atmosphere with velocities greater than some 3000 km/sec. FAN and SCHULTE (1954) have reported that in aurorae observed from Yerkes Observatory the Balmer lines are always present in auroral arcs but absent in auroral rays. From these observations, CHAMBERLAIN (1954b) concludes that protons could be the cause of auroral arcs, since the calculated hydrogen luminosity curves are in reasonable agreement with the observed luminosity curve of arcs. In another paper, CHAMBERLAIN (1955) shows that it might be possible for the protons to provide the necessary acceleration of ionospheric electrons, the electron current becoming constricted along narrow paths which appear as rays.

Other works have considered the occurrence and intensity of hydrogen in aurorae on a statistical basis in an attempt to correlate the emissions with magnetic activity and its latitude distribution (MONTALBETTI and VALLANCE JONES, 1957; VEGARD, 1958; GARTLEIN, 1952). MONTALBETTI and VALLANCE JONES found a negative correlation between H_α emissions and magnetic activity at Churchill whereas a positive correlation was observed at Saskatoon. GARTLEIN found no correlation for observations made at Ithaca, New York.

* The research reported in this paper was sponsored by the Radio Physics Laboratory, Defence Research Board, under contract.

† Radio Physics Laboratory, Defence Research Board, seconded to Physics Department, University of Saskatchewan, Saskatoon, Saskatchewan, Canada. Present address: Defence Research Northern Laboratory, Fort Churchill, Manitoba.

These deductions were based on observations made with spectrographs which require relatively long exposure times making a detailed study of the temporal variations of the emissions difficult. HUNTEN (1955), has pointed out from a number of preliminary photoelectric measurements that the Balmer line H_β shows sudden variations and that its occurrence cannot be predicted. For these reasons it seemed very desirable to make a further study of the hydrogen emissions using photoelectric techniques at a site well within the auroral zone.

This was done for the H_β emissions at Churchill, Manitoba, during the period, 10–24 February 1958, when a number of rockets instrumented for the measurement of electron and proton fluxes were fired into aurorae. The period was one with a large number of auroral displays, including the notable display of red aurora (Type A) on the night of 10–11 February. The results of these observations are reported in this paper.

2. INSTRUMENTATION

A scanning spectrometer, similar to that reported by HUNTEN (1953), except for the size of the grating (8 × 8 in.), was used to detect the H_β line in aurorae and airglow. The H_β line was chosen in preference to the stronger H_α line, because it is in a spectral region relatively free from auroral emissions and where the response of the 1P21 photomutiplier detector to incident radiation is large. The spectrometer was operated with a slit width of approximately 7 Å, scanning the region 4580 Å to 4950 Å in second-order in 12 sec sweeps every 15 sec. A secondary standard low brightness source (SHEPHERD, 1954) was used to calibrate the spectrometer so that absolute intensities could be determined.

As part of the International Geophysical Year programme other instruments were operating during this interval. These were a patrol spectrograph (Perkin–Elmer Corp., 1956), auroral intensity recorder (HUNTEN, 1956) and all-sky cameras (MEEK, 1955); and they provided additional information on night-sky conditions. A small grating spectrograph was operated most of the time to monitor the H_α line in the zenith. This instrument, described in a paper by MONTALBETTI and VALLANCE JONES (1957), was modified so that the spectroscopic plate moved past the camera lens at approximately 0·15 in./hr, and thus recorded the variation of the spectral features between 5577 Å and 6800 Å (H_α in particular) as a function of time. Because of the continuous motion of the plate during the entire night, only relatively strong emissions were detected.

3. OBSERVATIONS AND RESULTS

During the period of 10–24 February 1958, a relatively large number of auroral forms and types were observed. Type A aurora was viewed during the great display of 10 February and very strong active aurorae were seen on two separate nights.

The hydrogen emission was monitored continuously every night except during periods when heavy overcast was present. The results are given below according to the particular measurements that were undertaken. It should be noted that the programme was of an exploratory nature, and was not intended as a statistical study of any one point.

3.1. *Auroral forms and hydrogen emissions*

The spectral region 4580 Å to 4950 Å was scanned to provide intensities of the H_β line and of the (0, 2) and (1, 3) N_2^+ bands. In Table 1, the brightness of the hydrogen line, $B(H_\beta)$, and of the (0, 2) N_2^+ band, B(4709) expressed in Rayleighs (R), are tabulated in the third and fourth column respectively. The ratio of each pair of values is given in the fifth column, and is followed by a description of the aurora from which the value was obtained. Only the results for auroral forms brighter than intensity *I* are given in the table, because of the low signal to noise ratio for weak forms. There was, however, one occasion when strong hydrogen

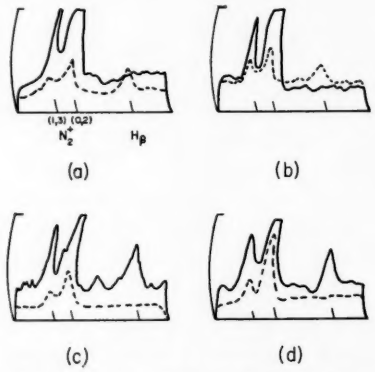


Fig. 1. Typical spectra taken over the period 10–24 February 1958, showing occurrence and non-occurrence of $H\beta$ with similar auroral forms viewed in different displays. (a) Homogeneous arcs; (b) bands; (c) rays; (d) diffuse surfaces.

emissions were recorded without the occurrence of visible aurora. This is included in the table (16–17 February), and is discussed at greater length in Section 3.2.

It is probable that the radiations from the homogeneous arcs and rays did not completely fill the field of view of the scanning spectrometer. Thus the brightness of the two spectral features may have been greater than that reported here. The ratio should still be correct.

An examination of these tabulated results indicate that the intensity of the hydrogen emissions do not correlate with specific auroral forms; at least for aurorae viewed from Churchill which is approximately at the centre of the auroral zone. This result is clearly illustrated in Figs. 1(a) to (d), where two typical spectra are reproduced for each of four auroral forms (homogeneous arcs, bands, rays and diffuse surfaces) observed at different times. In Fig. 1(c), for example, the top curve shows a spectrum obtained while the spectrometer was pointed at an isolated ray, while the bottom curve represents the spectrum secured at another time for a similar form. In the first case the H_β line is very strong while in the bottom curve there is no evidence of the hydrogen line. The situation was the same for the other three forms. It is estimated that a brightness of less than some 50 R would, in general, not have been detected with the spectrometer as it was employed. The fact that on a number of occasions H_β was detected in rays and not in homogeneous arcs is not in agreement with the results of FAN and SCHULTE (1954). However, in agreement with their results, no hydrogen emission was detected in

Table 1.

Date	Time CST	B(H_β) (R)	B(4709) (R)	H_β 4709	Auroral forms
10–11 February 1958	2232	260	370	0·70	RB type A
	2237	240	640	0·37	DS type A
	2255	340	850	0·40	DS no red
	2309	850	2900	0.29	DS type A
	2317	530	1800	0·30	DS no red
	0010	160	320	0·50	DS no red
	0024	850	1600	0·54	DS type A
	0045	*n*	3000	<0·02	C type A
	0054	*n*	3000	<0·02	B red and green
	0057	*n*	2300	<0·02	DS red and green
	0110	*n*	450	<0·1	DS type A
	0122	*n*	750	<0·07	DS type A
12–13 February 1958	2358	*n*	1000	<0·05	B II
	0040	*n*	5000	<0·01	Flaming
14–15 February 1958	2230	*n*	1800	<0·03	HA
	2310	*n*	—	—	HA
	0100	*n*	—	—	HA
15–16 February 1958	2350	*n*	7000	<0·007	Flaming
	0002	*n*	2400	<0·02	C
	0021	270	300	0·90	DS (zenith)
	0025	180	420	0·43	Rays
	0040	150	300	0·50	HA
	0130	*n*	900	<0·05	HA
	0120	210	270	0·78	B
16–17 February 1958	1950	550	*n*	—	none (zenith)
	2037	220	*n*	—	none (E. elevation 30°)
	2041	270	*n*	—	none (zenith)
	0024	300	210	1·4	Isolated ray
	0215	560	1800	0·31	RA and RB
	0247	250	230	1·1	RB
	0308	*n*	2000	<0·02	HA II
	0315	*n*	760	<0·07	B I
17–18 February 1958	1932	*n*	1600	<0·03	RB
	2000	*n*	1700	<0·03	RA
	2139	*n*	1000	<0·05	HA
	2253	*n*	760	<0·07	DS
	2302	*n*	2400	<0·02	HA II
	2312	*n*	820	<0·06	HB
	2314	*n*	630	<0·08	HA
	2350	370	570	0·65	RA type B
	0011	*n*	880	<0·06	C
	0035	*n*	2400	<0·02	B II

VOL. 14 1959

Table 1.—(*continued*)

Date	Time CST	B(H_β) (R)	B(4709) (R)	H_β 4709	Auroral forms
17–18 February 1958	0055	*n*	1500	<0·03	B
	0255	*n*	1100	<0·05	HA
	0257	*n*	350	<0·1	HB I
	0412	*n*	2500	<0·02	RB II
	0430	*n*	800	<0·06	DS
18–19 February 1958	2203	*n*	1000	<0·05	HA
	0028	*n*	1000	<0·05	RB
	0032	*n*	1000	<0·05	HB
20–21 February 1958	2109	*n*	6000	<0·008	RB II–III
	2117	*n*	550	<0·09	RB
	2256	320	690	0·46	HA
	2320	330	900	0·37	Faint cloud
	2331	270	520	0·52	Faint cloud
	2342	320	420	0·76	Faint cloud
	0005	160	370	0·43	R I
	0235	220	240	0·92	Cloudy
21–22 February 1958	1948	400	600	0·67	HA
	1952	*n*	2000	<0·02	B (zenith)
	2010	150	100	1·5	DS
	2049	200	60	3·3	HA
	2130	*n*	400	<0·1	HA and DS
	2300	*n*	1100	<0·05	R
	2302	170	470	0·35	R
	2336	170	330	0·51	B
	2331	*n*	4500	<0·01	RB
	2356	400	1600	0·25	B

n = not detected (<50 R)
R = unit of brightness in rayleighs.

the very strong active auroral forms such as very bright rayed arcs and rayed bands. This is also in agreement with the results of MONTALBETTI and VALLANCE JONES (1957) who observed no H_α emission in some twelve major displays at Churchill during the 1954–1956 period. In fact, the observations reported in Table 1 shows that, with the exception of type A aurora, no H_β emission was detected in aurorae when the (0, 2) N_2^+ band was brighter than 2000 R. This suggests that many auroral displays take place without the emission.

During the time of the great auroral display of 10–11 February when type A aurora was observed for some 3 hr, the hydrogen emissions varied considerably. At the beginning of the display, H_β appeared very strong in all the type A forms but in the later stages of the display it was not detected in either the more usual or the type A forms. This detailed study of the variation of the hydrogen emissions with auroral forms suggests that the impinging proton stream is not the direct cause of *all* auroral displays. While many forms were accompanied with hydrogen

emissions, which points to auroral excitation by protons, many of the same forms showed no such emissions. It is possible that during these times the proton stream was far south of the Churchill station and that the electron stream produced the auroral excitations. If this were so, the latitude displacement of the stream would have to vary rapidly since forms were observed within a 5 min interval in which the H_β line varied from an undetectable level to 170 R. (See results in Table 1 for the night of 21–22 February).

Taking CHAMBERLAIN's (1954a) deduction that 11 quanta of H_β radiation are emitted per proton on passage through the atmosphere a proton flux of 10^7–10^8 protons/cm^2 sec. is necessary to explain the H_β brightnesses from 150 to 850 R. The proton flux may have been larger than this, since the radiation from homogeneous arcs and rays probably did not fill the field of view of the spectrometer. However, the value is in good agreement with HUNTEN's (1955) observations.

3.2. *Hydrogen emissions in enhanced airglow*

Although most of the hydrogen emissions were detected during periods of visible aurorae, there was a $1\frac{1}{2}$ hr interval during the night of 16–17 February when strong H_β radiation was observed with no visible aurorae present. In Fig. 2 the brightness of the line, as measured in the zenith, is plotted against time (CST). No observations are available prior to 1950 hours owing to the spectrometer being employed on another programme. The hydrogen emission decreased steadily from a value of 500 R at the start of the observations to an undetectable level at approximately 2115 hours. A typical spectra obtained at the beginning of the run is shown in Fig. 3(a). It is clear that no (0, 2) N_2^+ radiation was detected, whereas the H_β line was very intense.

VOL. 14 1959

In Fig. 2 the relative intensity of the 5577 Å OI line, as measured in the zenith is shown for the same $1\frac{1}{2}$ hr interval. This information was taken from the auroral intensity recorder which scans a meridional strip of the sky 15° wide from horizon to horizon twice every 5 min (HUNTEN, 1956). The intensity of the 5577 Å line decreased continuously reaching a normal night-sky value at approximately the same time that the H_β emissions reached an undetectable level. The similarity of these two curves indicates an enhancement of night-sky radiations by the charge exchange reaction between the incoming protons and the atmospheric atomic oxygen.

All of the observations were made in the zenith except for a 5 min interval when the spectrometer was pointed in other directions. In Fig. 3(b) a typical spectrum obtained with the instrument pointing eastward at an elevation of 30° is shown (curve a), while curve (b) is the spectrum observed in the zenith some 30 sec earlier. Since the hydrogen emissions were approximately the same intensity in all directions, the proton stream must have covered a considerable area and it is unlikely that it was associated with any auroral form below the threshold of visibility.

No further H_β emissions were observed from 2115 hours until about 0024 hours when a weak auroral display developed and the hydrogen line reappeared in some, but not all, of the auroral forms that ensued. It was not observed again during intervals of no visible aurorae.

The small moving plate spectrograph, mentioned in Section 2, yielded similar qualitative results. The OI lines, 5577 Å, 6300 Å and 6364 Å, as viewed in the zenith, were enhanced at the start of the exposure and steadily decreased to a constant intensity at about 2130 hours. The H_α line followed the same behaviour except that it reached an undetectable level at 2120 hours, and was not recorded again on the spectroscopic plate during the remainder of the night.

The possibility that these enhanced radiations were associated with a major auroral display south of the station was considered. The all-sky camera film from Bird, an IGY station 250 km south of Churchill, and the auroral intensity recorder tape from The Pas, some 700 km south-west of Churchill, revealed that only very

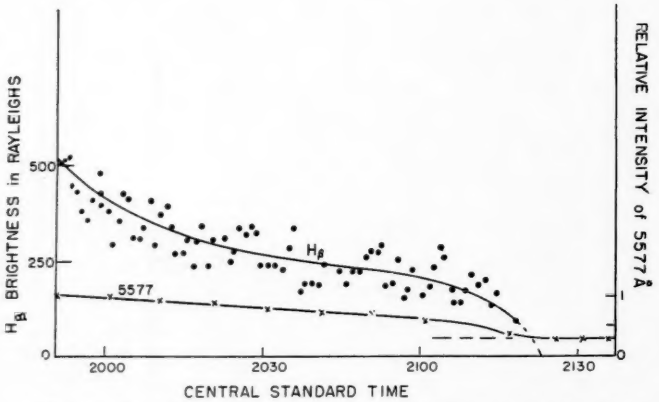


Fig. 2. Variation of the brightness of the $H\beta$ emissions and of the 5577 Å radiation from the zenith as a function of time during the night of 16–17 February 1958.

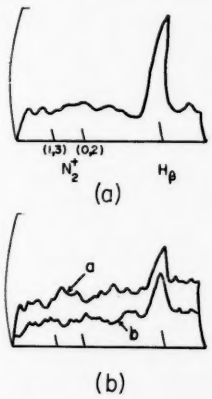


Fig. 3. Typical spectra showing the occurrence of $H\beta$ with no visible aurorae. (a) Taken at the start of observations, 1950 hours 16 February 1958. (b) Spectra showing brightness of $H\beta$ in two directions. Curve (a) is for 30° elevation in the east and curve (b) is for the zenith.

weak aurorae were present during the period under consideration. In addition, these aurorae were recorded only to the south of both stations. The evidence is substantial that the hydrogen emissions were not associated with an intense auroral display and that they did not precede one.

An attempt was made to determine the energy of the incoming protons by examining the profiles of the H_α line, obtained with the Churchill patrol spectrograph, and of the H_β line, observed with the scanning spectrometer. The details of the measurements are discussed in the next section. It is sufficient to point out here that the extent of the violet wing on both profiles indicated that the initial proton velocities did not exceed approximately 1200 km/sec. Such low velocity protons would not penetrate below the 130 km level. This is consistent with the observation that the green and red lines of OI were enhanced, while the (0, 2) N_2^+ band could not be detected.

The proton flux giving rise to the observed brightness of the H_β line during the

interval of strong hydrogen emissions may be calculated from CHAMBERLAIN's work (1957). Taking his published curve for the relative number of quanta emitted per unit diminution of velocity, and the result that a proton with an initial velocity greater than 4000 km/sec emits about 11 quanta of H_β radiation in its passage through the atmosphere, it is readily shown that a proton with an initial velocity of 1200 km/sec should emit about 1 quantum. The observed brightness of the H_β line at the beginning of the observations was 500 R, of which about one-third was probably the real zenithal brightness (see next section). This corresponds to a proton flux of the order of 10^8 protons/cm^2 sec.

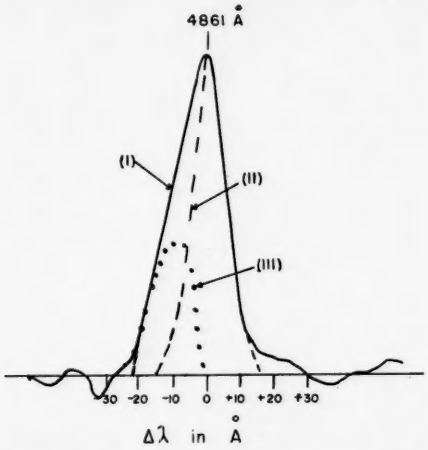


Fig. 4. Zenithal H_β line profile secured with scanning spectrometer on night of strong hydrogen emission without visible aurorae: (I) is observed profile; (II) is the assumed curve due to light scattering; (III) is the deduced true zenithal profile.

A continuum was observed on the spectra secured with the scanning spectrometer, as is evident from the examples shown in Figs. 3(a) and 3(b). Although it was also observed during periods of aurorae, it was more clearly defined during the period of the strong hydrogen emissions, and appeared more intense toward the horizon than the zenith (compare (a) and (b) in Fig. 3(b)). The occurrence of this continuum has been reported by OMHOLT (1957) for those stages of an auroral display when rays or pulsating forms occur. The present observations point to its more frequent occurrence. Further measurements are required to establish the intensity of this continuum, since the present results may be contaminated by scattered light from ground level.

3.3. *Hydrogen Balmer line profiles*

During the time of strong hydrogen emissions with no visible aurorae, the spectrometer scanned a comparatively large spectral region; and hence the profile of the H_β line could only be measured to a limited accuracy. An example of such a profile from a zenithal observation at the beginning of the period is shown in Fig. 4, and is labelled (1). To within about ± 2 Å, the peak of the profile is unshifted and the violet wing extends some 23 Å from the peak. If the profile is due entirely to protons travelling along the line-of-sight, its shape can only be reconciled with a

velocity distribution in the stream, most of the protons having extremely low initial velocities (<100 km/sec). Since obviously the incoming protons covered a large extent of sky a more plausible explanation appears to be that the scattered light from all regions seriously modified the true zenithal profile. Assuming that the horizon profile is symmetrical, (as MEINEL's (1951) observations on aurorae have shown), and that the red wing of the profile is due entirely to the scattered light from the horizon, a symmetrical profile was traced out in Fig. 4 (curve II). The

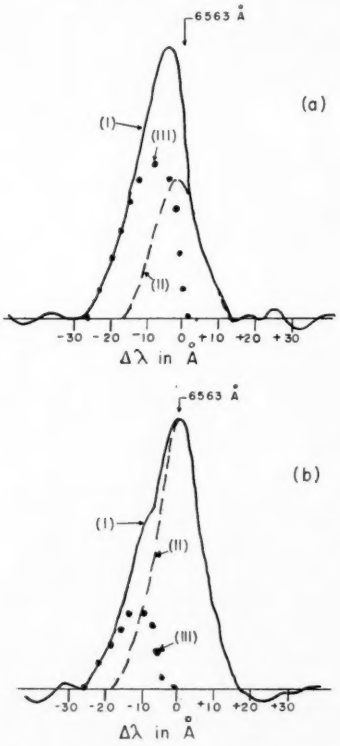


Fig. 5. H_α line profiles taken with patrol spectrograph on the night of 16–17 February 1958. (a) Horizon profile. (b) Zenithal profile. For explanation of various curves see text.

difference between these two profiles is then the actual zenithal profile, and corresponds approximately to what would be observed if all the protons entered the upper atmosphere with a velocity of 1200 km/sec. The actual sharp cut off on the violet side of the H_β line due to monoenergetic protons would be smoothed out to a linearly rising front as the exit slit scanned across the spectral line. The peak would also occur about 8 Å away from the extreme end of the wing. It is clear that the actual brightness of the H_β line as observed in the zenith would be about one-third of the measured brightness.

In order to pursue the above considerations further, the spectra secured with the Churchill patrol spectrograph were analysed. The spectra resulted from a single exposure covering the interval from 1955 hour to 2255 hours and included the period of strong hydrogen emissions. The H_α profile for both zenith and horizon could be measured since a meridional strip of the sky is focussed on the entrance

slit. In Fig. 5 (a), the profile measured in the zenith is shown (curve I), while the horizon profile is reproduced in Fig. 5 (b) (curve I). The two curves were normalized to equal intensity at the peak. It is evident that both profiles are asymmetrical with the horizon profile having the greatest asymmetry, that the peak

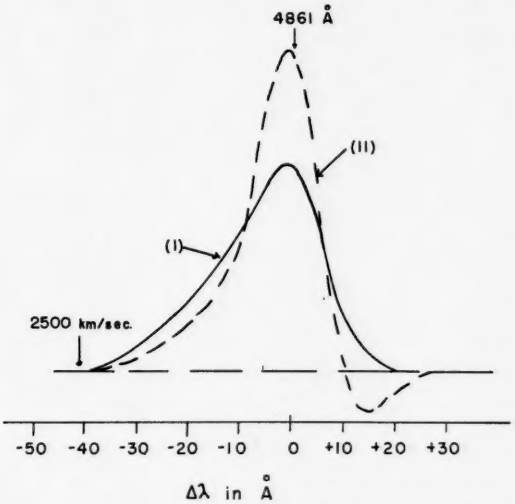


Fig. 6. H_β line profile secured during night of 20–21 February 1958, with scanning spectrometer viewing weak zenithal aurorae. Curve (I) is observed profile while (II) is profile corrected for instrumental broadening.

of the zenithal profile appears displaced some 4 Å, and that, to the accuracy of the measurements, the horizon profile is unshifted. Continuing on the assumption that the horizon profile should be symmetrical, curve (II) of Fig. 5(b) was assumed as the true horizon profile and was subtracted from the observed one. The resultant curve (III) should then correspond closely to the zenithal radiation that was scattered by the lower atmosphere.

The zenithal profile was examined with the view that the red wing could be due to scattered horizon light. A symmetrical and undisplaced horizon profile was matched to the wing, and subtracted from the measured profile to yield the supposed true zenithal profile (curve III) of Fig. 5(a)). (There is reasonable agreement between this curve and that deduced from the horizon profile.) In this analysis, it turns out that the amount of horizon radiation that is scattered in the zenith is greater than the reverse, which would be expected since the hydrogen emissions appeared to originate from all directions. Furthermore it should be noted by comparing curve (III) of Fig. 4 to curve (III) of Fig. 5(a) that the scattered light contributed less to the H_α than to the H_β profiles as measured in the zenith. This would be expected from the wavelength dependence of rayleigh scattering.

The violet wing of about 25 Å indicates a line-of-sight proton velocity of 1100 km/sec in agreement with the spectometer measurement of the H_β profile. However, the shape of the H_α line profile does not agree too well with what would be expected if all the protons were monoenergetic, and had an initial velocity of 1100

km/sec. This may not be serious since the exposure covered a relatively long period of time during which the initial proton velocity could have decreased.

The results from both the scanning spectrometer and from the patrol spectrograph may be qualitatively explained if light scattering is assumed to play a major role in the measurements. The above analysis can only be approximate since the curve which was assumed to be due to scattering could be considerably modified by contributions from scattered radiations from all regions of the sky. The accuracy of the measurements do not warrant the more detailed analysis.

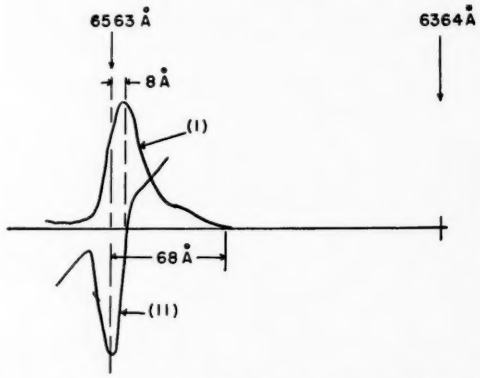


Fig. 7. H_α line profile from moving plate spectrograph pointing in zenith. (I) is observed profile of emissions associated with aurorae; (II) is H_α Fraunhofer line in moonlight.

An accurate measurement of the H_β line profile during periods when the hydrogen emissions were associated with aurorae was attempted on the night of 20–21 February 1958. The scanning spectrometer was employed with a 7 Å slit width and adjusted to scan a relatively small spectral region of 100 Å. A series of six scans, all taken on rather weak zenithal aurora during a time interval of some 2 min were averaged and yielded the profile shown in Fig. 6 curve (I). This was corrected for instrumental broadening according to the graphical method reported by BRACEWELL (1955) and is shown as curve (II) of the same figure. The observed shift of about 1 Å and a violet wing extending some 38 Å from an undisplaced peak agrees very well with VEGARD's (1956) results of a 1·5 Å and a 35 Å wing.

The H_α profile was measured on a spectroscopic plate which was exposed during the night of 25–26 February in the moving-plate spectrograph. The instrument was pointed at the zenith for the entire night, and showed strong hydrogen emissions associated with zenithal aurorae. On other occasions when the H_α line was detected with the spectrograph, the plate density was usually too low for good microphotometering. In Fig. 7, curve (I) corresponds to the observed profile (no correction was made for plate density vs. light intensity, or for instrumental broadening), while curve (II) is the Fraunhofer H_α line observed in moonlight. A violet wing of some 68 Å and a shift in the peak of 8 Å is observed which agrees very well with MEINEL's (1951) observations of a 70 Å wing and a 10 Å shift.

Although the H_α and H_β line profiles were secured on different nights, they agree with independent measurements made by other workers. The violet wing of

VOL
14
195

the H_α line indicates an initial proton velocity equal to or greater than 3100 km/sec while the H_β wing gives an initial velocity of about 2300 km/sec. It should be stressed, however, that the shift of the H_β line is far less than expected. The cross-section for the emission of H_β radiation is a maximum for protons having a velocity of about 2270 km/sec while in the case of the H_α emission this occurs at about 2000 km/sec (CHAMBERLAIN, 1958). It is possible that scattering by the lower atmosphere plays an important role in modifying the true H_β zenithal profile. Further measurements are planned to investigate this point.

4. CONCLUSIONS

The results of this preliminary investigation may be summarized as follows:

(1) A systematic relationship does not appear to exist between various auroral forms and hydrogen emissions at places within the auroral zone. Strong displays occur without detectable hydrogen emissions. The intensity of the hydrogen line varied rapidly and its occurrence could not be predicted. The proton flux associated with displays showing strong H_β emissions was calculated to be the order of 10^7–10^8 protons/cm^2 sec.

(2) Hydrogen emissions are not necessarily associated with visible aurorae, since strong H_β radiations were observed over a relatively long period without visible aurorae. On this occasion the incoming protons covered a large area of the sky. Also they probably were monoenergetic and had an initial velocity of 1100 km/sec, although this latter point is by no means well established. A proton flux of some 10^8 protons/cm^2 sec was deduced from the initial brightness of the H_β line.

(3) The measurements of the line profiles during the period of strong hydrogen emissions without visible aurorae can be explained by assuming that scattering of light plays an important role in modifying the true profiles. If this is the case, correction for scattering should also be considered in profile measurements on hydrogen lines associated with aurorae.

(4) Continuous photoelectric detection of hydrogen lines is not only possible but very desirable. The observations also point to a need for a more thorough investigation of the occurrence of the hydrogen emissions associated with aurorae and airglow. Accurate photoelectric measurements of the line profiles should yield important information on the variation of proton velocities and establish the role of light scattering to the measurements.

Acknowledgments—The author wishes to acknowledge the assistance of Mr. D. McEWEN and Mr. H. LUTZ in making the observations at Churchill. Sincere thanks are extended to Professors B. W. CURRIE and D. M. HUNTEN, Physics Department, University of Saskatchewan, the former for his interest and encouragement in the programme, and the latter for valuable discussions and suggestions concerning the analysis of the data. It is worth noting that the research was only possible through a co-operative effort between the Physics Department of the University of Saskatchewan and the Defence Research Northern Laboratory.

REFERENCES

BRACEWELL R. N. 1955 *J. Opt. Soc. Amer.* **45,** 873.
CHAMBERLAIN J. W. 1954a *Astrophys. J.* **120,** 360.

VOL. 14 1959

CHAMBERLAIN J. W.	1954b	*Astrophys. J.* **120,** 566.
CHAMBERLAIN J. W.	1955	*The Airglow and the Aurorae* (Edited by ARMSTRONG E. and DALGARNO A.). Pergamon Press, London.
CHAMBERLAIN J. W.	1957	*Astrophys. J.* **126,** 245.
CHAMBERLAIN J. W.	1958	*Advances in Geophysics* (Edited by LANDSBERG H. and VAN MIEGHEM J.). Academic Press, New York.
FAN C. Y. and SCHULTE D. H.	1954	*Astrophys. J.* **120,** 563.
GARTLEIN C. W.	1952	*Mem. Soc. Sci. Liège* **12,** No. 1–2, 195.
HUNTEN D. M.	1953	*Canad. J. Phys.* **31,** 681.
HUNTEN D. M.	1955	*J. Atmosph. Terr. Phys.* **7,** 141.
HUNTEN D. M.	1956	*J. Opt. Soc. Amer.* **46,** 578.
MEINEL A. B.	1951	*Astrophys. J.* **113,** 50.
MEEK J. H.	1955	*All-Sky Camera.* University of Saskatchewan Report.
MONTALBETTI R. and VALLANCE JONES A.	1957	*J. Atmosph. Terr. Phys.* **11,** 43.
OMHOLT A.	1957	*Astrophys. J.* **126,** 61.
Perkin-Elmer corporation, Norfolk, Conn.	1956	*Instruction Manual for the Model* 173 *I.G.Y. Auroral Spectrograph.*
SHEPHERD G. G.	1954	*Scientific Report* AR-16, *Contract AF* 19(122)–152. University of Saskatchwan.
VEGARD L.	1939	*Nature, Lond.* **114,** 1089.
VEGARD L.	1956	*Geofys. Publ.* **19,** No. 9.
VEGARD L.	1958	*Geofys. Publ.* **20,** No. 4.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 213 to 228. Pergamon Press Ltd. Printed in Northern Ireland

Some indices of solar activity based on ionospheric and radio noise measurements

C. M. MINNIS and G. H. BAZZARD
(Official Communication from the D.S.I.R. Radio Research Station, Slough)

(*Received* 9 *July* 1958)

Abstract—The critical frequencies of the *E*- and *F2*-layers of the ionosphere are closely controlled by the level of solar activity but, at any fixed point on the earth's surface, they also vary with season owing to the annual change in the solar zenith angle. A critical frequency can only be used to provide an index of solar activity if such seasonal variations can be eliminated or reduced to very small proportions. Methods of achieving this are described and monthly mean values of an *E*- and an *F2*-layer index are tabulated for the period 1938–1957. The correlation of these indices with each other and with the solar noise flux at $\lambda 10{\cdot}7$ cm is high. Some possible practical applications of such indices are discussed briefly.

VOL. 14 1959

TERMINOLOGY

A_E, A_{F2}	The ratios $f_E(R)/f_E(0)$ and $f_{F2}(R)/f_{F2}(0)$.
a_m, b_m	Constants for the month m.
Ch_E	$f_E{}^4 \sec \chi$.
f_E, f_{F2}	The critical frequencies of the E and $F2$-layers.
$f(R)$, $f(0)$	Critical frequencies for sunspot numbers R and 0.
f_m, f_n	Critical frequencies for consecutive months m and n.
g_m, h_m	Constants for the month m.
g_0, h_0	Constants applicable to all months.
I_E	$\lambda(m, \Phi)Ch_E$
I_{F2}	The mean value of R' for several observatories.
$I_{F2}(E)$	An index analogous to I_{F2} but based on Ch_E instead of R.
k	Constant assumed to be independent of m and R.
$k(m,R)$	k represented as a function of m and R.
m, n	Symbols indicating a month m and the following month n.
p_{mn}, q_{mn}	Constants relating to consecutive months m and n.
R	The monthly mean relative sunspot number.
R'	$(f_{F2} - a_m)/b_m$
S	The flux of solar ionizing radiation.
S_R, S_0	The values of S at sunspot numbers R and 0.
Φ	The solar radio noise flux at $\lambda 10{\cdot}7$ cm.
$\lambda(m, \Phi)$	A function which varies with season and solar activity.
χ	The solar zenith angle.
ρ	f_m/f_n. Theoretical value: $(\sec \chi_n/\sec \chi_m)^{1/4}$.

1. TYPES OF INDEX

OBSERVATIONS of many solar and terrestrial phenomena give a clear indication of the existence of a persistent 10- or 11-year cycle in the level of solar activity and it is often necessary to express this level in terms of a numerical index number. In any discussion of such indices it is important to distinguish between primary

indices, which may be regarded as direct measurements of solar activity, and secondary indices, which are derived from data relating to terrestrial phenomena, the variations in which also contain marked seasonal components which must be eliminated.

1.1. *Primary indices*

This type of index is based on the direct measurement of some characteristic of the sun itself such as the areas of calcium faculae or sunspots, the power flux of solar radio noise, or the relative sunspot number. It is assumed that the magnitude of an index of this kind is not controlled by any terrestrial influence and that it depends only on conditions on the sun. Apart from sunspot numbers, the only primary index discussed in this paper is the power flux of solar radio noise at a wavelength of 10·7 cm.

1.2. *Secondary indices*

Solar activity controls many characteristics of the earth's upper atmosphere and, in consequence, several ionospheric and geomagnetic parameters could, in principle, be used as measures of solar activity. However, such parameters nearly always vary with time of day and with season. If it is required to eliminate these unwanted components, so as to isolate the component which is attributable to changes in solar activity, this may be achieved either by empirical methods, or by calculations based on some theory of the relation between solar activity and the quantity which is measured.

Indices from which large diurnal and seasonal components have been eliminated will be referred to as secondary indices. Except for the calculated indices Ch_E and I_E, the secondary indices described in this paper have been obtained by using empirical methods to eliminate seasonal effects from the basic data.

2. Applications of Solar Indices

The term "solar activity" has no precisely defined general meaning. When an index of solar activity is required for a particular purpose, it is essential first to ask what is implied by the term and then to adopt the index which seems best suited to the problem under investigation. The application of such indices to several different problems is discussed briefly in the remainder of this section.

2.1. *Investigation of solar–terrestrial relations*

As already mentioned, many geophysical phenomena are closely related to changes in solar activity. The active agent which is responsible for the existence of the ionosphere and for many geomagnetic effects is the X-ray and ultra-violet radiation emitted by the sun. In studies of the ionosphere and geomagnetism, it is often necessary to correlate observed phenomena with estimates of the changes in the intensity of the ionizing radiation. The ideal index would obviously be a direct measure of the radiation flux but, owing to atmospheric absorption, this radiation can not be detected except during occasional flights of high-altitude rockets or earth satellites. In principle, however, it may still be possible either to calculate the flux indirectly or to derive an index which is proportional to it from ionospheric or geomagnetic measurements.

It is by no means certain that it is sufficient to consider only one kind of ionizing radiation; several different indices may be required ultimately for different applications.

2.2. *Medium-term forecasts*

In the field of radio-communications, forecasts of future conditions in the ionosphere are often required about 6 months ahead. These forecasts are based on the extrapolation of the current trend in solar activity. For such extrapolations, the flux of ionizing radiation, or an index proportional to it, would be ideal and preferable to the relative sunspot number which, in spite of its erratic behaviour, is the index normally used at present. However, a less stringent specification of an index for this purpose is possible: the main requirement is for an index which is closely correlated with the ionospheric parameter of which forecasts are required and it does not matter if the index is not linearly related to the flux of ionizing radiation.

2.3. *Long-term forecasts*

Forecasts of solar activity for periods more than 1 or 2 years ahead can not be made accurately. Although monthly indices are needed for the applications referred to in Sections 2.1 and 2.2, a 12-month running mean of some index is normally adequate for long-term forecasts since these are usually made by fitting an empirical relation to the current cycle of the smoothed sunspot number. Hence even a very erratic monthly index, such as the sunspot number, is adequate since the 12-month average smooths out the more violent fluctuations. Further, the sunspot number is the only index for which a series of values extending over a long period is available; in making long-term forecasts, the availability of such a series is much more important than it is when making medium-term forecasts.

VOL. 14 1959

2.4. *Short-term forecasts*

The preparation of forecasts of ionospheric and geomagnetic disturbances up to about 1 month ahead is sometimes attempted. Such forecasts are based mainly on a study of the life history and the detailed development of individual local centres of solar activity; the medium- and long-term changes in the general level of activity are only of secondary importance in this context. In consequence, indices of the type under discussion in this paper are not appropriate to short-term forecasts.

2.5. *The required indices*

From the preceding paragraphs it may be concluded that long-term forecasts of solar activity must continue to be based on sunspot numbers and that short-term forecasts require a detailed knowledge of the active solar regions present at any time and their life histories.

There remains, however, a need for indices of solar activity which might be used in making medium-term forecasts and in the study of terrestrial phenomena which are influenced strongly by the intensity of the solar ionizing radiation. These requirements could be met by deriving indices either from ionospheric data or from measurements of the solar noise flux at a frequency which is high enough

to ensure that the radiation originates in the lower, rather than the upper, levels of the corona. The derivation of such indices and the relations between them are discussed in Sections 3 and 4.

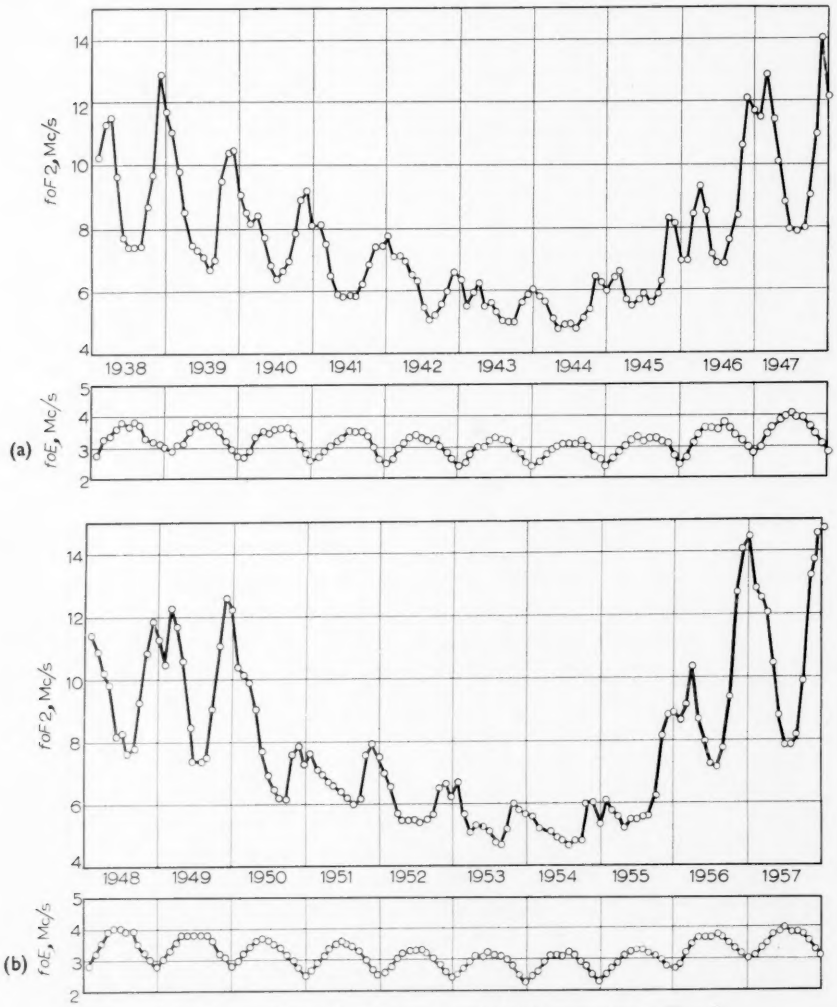


Fig. 1 (a and b). Monthly mean critical frequencies at Slough.

3. Methods for Deriving Secondary Ionospheric Indices

If, for a given place and time of day, the critical frequency, f, of an ionospheric layer is plotted against time, curves such as those in Figs. 1(a) and 1(b) will result. These curves obviously include a long-term change, which is due to the solar cycle, and a prominent seasonal component due to the rotation of the earth round the sun. Thus, for a month m when the flux of ionizing radiation is S, we may write:

$$f = \psi(S, m) \tag{1}$$

The derivation of a solar index based on measured values of f implies that it is possible to discover the nature of the function $\psi(S, m)$ and to express equation (1) in the form

$$S = \xi(f, m) \tag{2}$$

The forms of equations (1) and (2) for the *E*- and *F2*-layers will be discussed in Sections 3.1 and 3.2.

3.1. *E-layer; a calculated index*

The seasonal variations in f_E for a fixed time of day can be accounted for to a fairly high degree of accuracy by assuming that Chapman's theory is applicable to the *E*-layer. With this assumption, it follows that

$$f_E = (kS \cos \chi)^{1/4} \tag{3}$$

and

$$kS = f_E{}^4 \sec \chi = Ch_E \tag{4}$$

If k remains constant the *E*-layer "character figure" (Ch_E) ought evidently to be proportional to the intensity of the ionizing radiation (APPLETON and NAISMITH, 1939).

It has been found that, in practice, the variations in Ch_E still contain a small seasonal component. To compensate for this, it is necessary to multiply Ch_E by a factor which is a function of season and solar activity; the resulting index (I_E), defined by equation (5), is tabulated in Appendix II and explained in more detail in Appendix I B:

$$I_E = \lambda(m, \Phi)Ch_E \tag{5}$$

3.2. *F2-layer*

3.2.1. *An empirical index related to R.* For the *F2*-layer, the form of equation (1) is not known and hence relations between f_{F2} and S, such as equations (2) and (3) are not available. For this reason, indices of solar activity based on *F2*-layer measurements must be derived by some empirical method. Such methods were originated by GOODALL (1939) and have since been used with modifications by PHILLIPS (1947), APPLETON and PIGGOTT (1954) and MINNIS (1955).

For any given location, month and time of day, the monthly mean critical frequency of the *F2*-layer can be expressed approximately, as a function of sunspot number, by a linear relation:

$$f_{F2} = a_m + b_m R \tag{6}$$

Once a series of measurements of f_{F2} has been made at a given location during a complete solar cycle, it is not difficult to determine a_m, b_m for each month of the year. It is then possible to use these constants, together with the measured values of f_{F2}, to give a crude index of solar activity (R') defined by:

$$R' = (f_{F2} - a_m)/b_m \tag{7}$$

R' may be regarded as the component of R which is closely correlated to the mean intensity of the *F2*-layer ionizing radiation. In effect, the erratic fluctuations in R, which are often not related to any corresponding changes in the *F2*-layer, are smoothed out by using the ionization in the layer itself as an indicator of the

VOL. 14 1959

intensity of the incident radiation. The line defined by equation (6) may be regarded as a calibration, for a particular month, of the experimental values of f_{F2} against the sunspot number scale. Since the same scale is used for all months and since it is not subject to any seasonal changes, the variations in the index R' will not contain a seasonal component.

The magnitude of R' depends to some extent on where the measurements were made and on the degree of magnetic disturbance at the time. MINNIS has minimized this geomagnetic control by averaging the values of R' at three locations to give an index, I_{F2}, defined by:

$$I_{F2} = 1/3 \sum_{1}^{3} R' \tag{8}$$

This index is described more fully in Appendix I C and tabulated in Appendix III.

The relations between I_{F2} and the two indices I_E and Φ show that the relation between I_{F2} and S is approximately linear. Even though this relation is not precisely known, I_{F2} may still be used for purposes such as those mentioned in Section 2.2.

3.2.2. *An empirical index related to* Ch_E. A second *F2*-layer index, $I_{F2}(E)$, has been compiled which is exactly analogous to I_{F2} except that, in equations (6) and (7), f_{F2} is related to Ch_E instead of to R. This index is discussed later but the actual values have not been reproduced.

3.3. *E- and F2-layers: the relative critical frequency*

A possible alternative method of eliminating the seasonal component from a series of critical frequencies has been suggested by ALLEN (1948) who calculated the ratio (A) of the mean critical frequency, $f(R)$, for a given month at sunspot number R, to an extrapolated value, $f(0)$, for the same month at zero sunspot number. For the *E*- and *F2*-layers, equations (3) and (6) lead respectively to:

$$A_E = f_E(R)/f_E(0) = (S_R/S_0)^{1/4} \tag{9}$$

$$A_{F2} = f_{F2}(R)/f_{F2}(0) = 1 + b_m R/a_m \tag{10}$$

Thus, provided k in equation (3) remains constant, A_E ought to represent the ratio of the intensity of the ionizing radiation at any epoch to that at zero sunspot number. The characteristics of A_{F2} will depend on how the ratio b_m/a_m varies with season and on whether any such variations can be eliminated by combining the values of A_{F2} for several locations.

3.4. *The solar radio noise flux at* λ10·7 cm

Regular measurements of the incident radio noise flux of solar origin at λ10·7 cm have been made during the last decade at Ottawa and Tokyo. Only very small seasonal effects, caused by changes in the effective sky temperature, are present in the magnitude of the flux and, since appropriate corrections can be applied to eliminate these, such flux measurements may be regarded essentially as a primary index of solar activity. Revised values of the flux measured at Ottawa have been described by MEDD and COVINGTON (1958) and their monthly mean values, Φ, are tabulated in Appendix IV and plotted in Fig. 2 for comparison with I_E and I_{F2}.

VOL
14
195

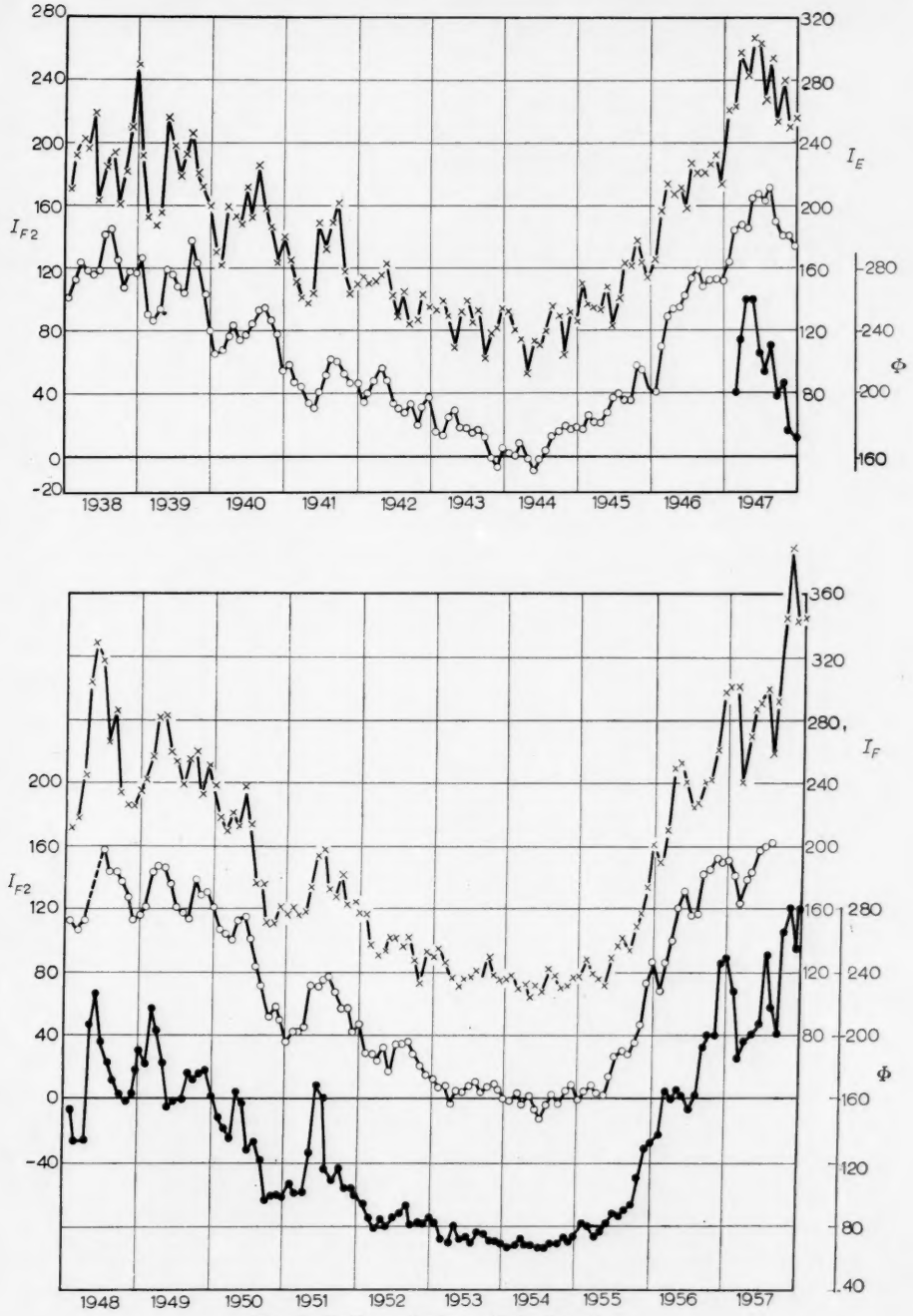
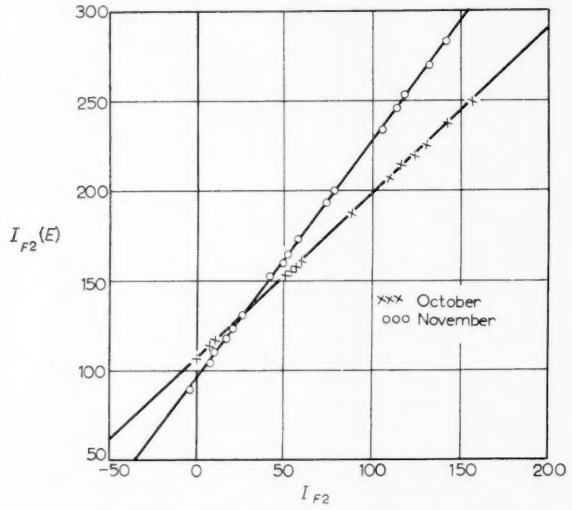


Fig. 2. Indices of solar activity based on:
× × × *E*-layer ionization (I_E);
○ ○ ○ *F2*-layer ionization (I_{F2});
● ● ● Radio noise at λ10·7 cm (Φ).

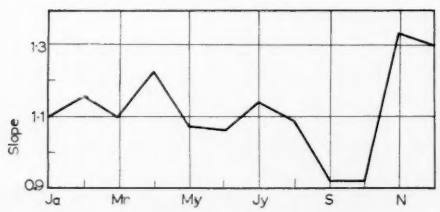
4. THE RELATIONS BETWEEN THE INDICES

4.1. I_{F2} and $I_{F2}(E)$

The indices I_{F2} and $I_{F2}(E)$ have been derived, in effect, by converting the measured values of f_{F2}, which vary with season, into equivalent values of R and Ch_E respectively, both of which are assumed to be independent of any seasonal effects. Hence it would be expected that the variations with time in I_{F2} and $I_{F2}(E)$ should not contain any seasonal component and that a single relation between these two indices should be valid throughout the year.



(a) The lines for October and November



(b) The annual variation in the slope

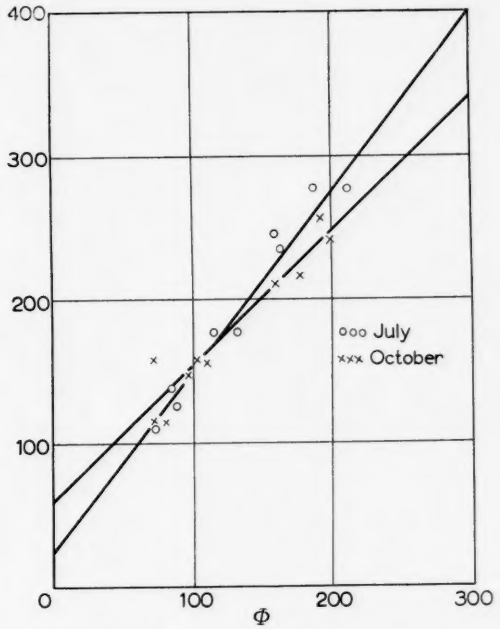
Fig. 3. Month–month changes in the slope of the linear relation between $I_{F2}(E)$ and I_{F2}.

In spite of these expectations, it has been found that for some months there are significant differences in these relations, two of which are shown graphically in Fig 3(a). The monthly values of the slope of the linear relations are shown in Fig. 3(b) and the existence of variations in slope suggests that, in actual fact, either R or Ch_E must, contrary to the earlier assumption, be subject to seasonal changes in magnitude.

VOL.
14
1959

4.2. Φ *and* Ch_E

As reported by MINNIS and BAZZARD (1958), the correlation coefficient between pairs of values of Φ and Ch_E is greater when each month is considered separately than when the data for all months are treated together. This is accounted for by the month to month variation in the relation between Φ and Ch_E; two such relations are shown in Fig. 4(a) and the variation in slope during the year in Fig. 4(b).



(a) The lines for July and October

(b) The annual variation in the slope

Fig. 4. Month–month changes in the slope of the linear relation between Ch_E and Φ.

There is an obvious similarity between the curves of Figs. 3(b) and 4(b) and this suggests that in both cases the monthly variations are due to some variable factor which is common both to the relation between I_{F2} and $I_{F2}(E)$ and to that between Φ and Ch_E. This common factor is Ch_E which is used in the derivation of $I_{F2}(E)$, and it is concluded that, contrary to the implications of equation (4), Ch_E is not independent of seasonal effects. This conclusion is confirmed by the fact that there is no significant seasonal variation in the relation between Φ and I_{F2}.

4.3. *Ch_E (and f_E) for consecutive months*

Further confirmation of the existence of a seasonal component in Ch_E has been obtained without resorting to a comparison with some other index such as Φ or I_{F2}. From the definition of Ch_E in equation (4)

$$f_E{}^4 \sec \chi = Ch_E = kS \tag{4}$$

Hence if the level of solar activity, as indicated by S, remains constant from one month (m) to the next (n),

$$f_m{}^4 \sec \chi_m = f_n{}^4 \sec \chi_n \tag{11}$$

and

$$\rho = f_m/f_n = (\sec \chi_n/\sec \chi_m)^{1/4} \tag{12}$$

Using the monthly mean values of f_E at Slough for the period 1938–1957, the ratio f_m/f_n has been plotted against sunspot number for each consecutive pair of months. Three of these plots are reproduced in Fig. 5 together with the appropriate value of $(\sec \chi_n/\sec \chi_m)^{1/4}$. Even though the points are scattered owing to the variations in solar activity from one month to the next, it seems reasonable to conclude that the ratio f_m/f_n varies with both season and solar activity. A line through the points can not be represented accurately by equation (12) and the actual relation has the form

$$f_m/f_n = (\sec \chi_n/\sec \chi_m)^{1/4} + p_{mn}R + q_{mn} \tag{13}$$

A relation of this kind leads to the conclusion that the factor k in equations (3) and (4) is not constant, but that it varies both with season and with the level of solar activity. This appears to be the ultimate reason for the fact that the magnitude of Ch_E is not controlled solely by solar activity and why it is necessary to apply the correction referred to in Section 3.1 to obtain an index which is free from seasonal control.

Sample plots of the type shown in Fig. 5 have also been made for a few months for Watheroo and Port Stanley. These show discrepancies which are similar to, but not identical with, those for Slough and it seems clear that the seasonal departures of f_E from the values expected for a Chapman layer are not the same at all locations.

4.4. *A_E and I_E*

The implication of the preceding discussion is that k in equation (4) is not a constant but is a function, $k(m,R)$ of the month and solar activity. This affects the assumption on which the definition of A_E was based and equation (9) becomes

$$A_E = f_E(R)/f_E(0) = [k(m,R)S_R/k(m,0)S_0]^{1/4} \tag{14}$$

In general, the ratio $k(m,R)/k(m,0)$ will be a function of both m and R and, in consequence, there will be a seasonal variation in the magnitude of A_E quite apart from variations due to changes in S_R. The existence of the seasonal component in the values of A_E which have been tabulated by ALLEN has been confirmed by plotting A_E against the corresponding values of I_E. The resulting relations show that even if the radiation intensity as indicated by I_E remains constant, A_E varies with season over a range of about ± 10 per cent. An investigation of the relations

between A_E and I_E and between Ch_E and Φ might provide useful information on the form of the function $k(m,R)$ but this has not yet been attempted.

An additional difficulty associated with the use of both A_E and A_{F2} is that of extrapolating the measured critical frequencies to obtain $f_E(0)$ and $f_{F2}(0)$, the critical frequencies at zero sunspot number. The method adopted by ALLEN appears to lead to presumed values of $f_E(0)$ and $f_{F2}(0)$ which are too high in northern summer and this, in turn, implies that the values of A_E and A_{F2} which

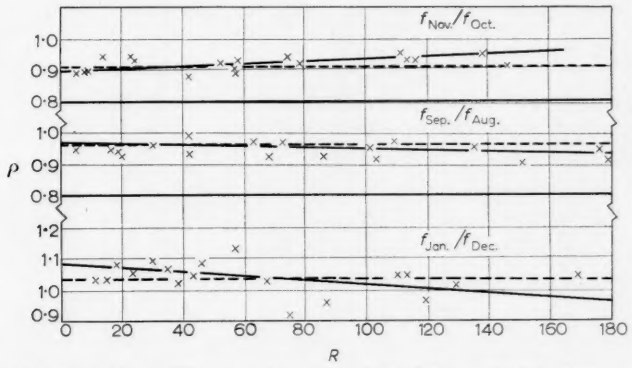


Fig. 5. $\times\times\times$ Variation with sunspot number (R) of the ratio (ρ) of the noon E-layer critical frequencies (f) for consecutive months (Slough 1938–1957). ——— Expected value of ρ.

he has tabulated contain an apparent seasonal component which is additional to that which arises from the function $k(m,R)$ for A_E, and from the ratio b_m/a_m for A_{F2}.

4.5. *A_{F2} and I_{F2}*

Plots of A_{F2} against I_{F2} disclose a seasonal variation in the relation between these two indices. The variations are such that the discrepancies are fairly small for $I_{F2} \simeq 50$, but are much greater for higher and lower values. The difficulty, referred to in Section 4.4, of finding the appropriate values of $f_E(0)$ and $f_{F2}(0)$ must be partly responsible for these variations. However, even if this source of error were eliminated, there remains the more serious problem of eliminating from A_{F2} the large seasonal variations which are characteristic of any one location. For example, for a constant level of solar activity represented by $I_{F2} = 150$, A_{F2} varies from 1·6 in northern summer to 2·6 in winter; at Huancayo the range is only 1·6–2·1, but the seasonal maxima and minima are less regular. By averaging A_{F2} for ten observatories, ALLEN has reduced, but not eliminated, the resultant seasonal variation, but it would still be quite large at levels of solar activity above and below that to which ALLEN's investigation refers.

4.6. *Numerical relations between Φ, I_{F2} and I_E*

Of the indices which have been discussed, only Φ, I_{F2} and I_E appear to be free from seasonal changes in magnitude and to be capable of representation, as a

VOL. 14 1959

function of one of the other two indices, by a single numerical relation valid for all months. Second order relations have been computed to represent the approximate relations between the three indices, equations (15) to (17), and from these the value of an index can be obtained from the measured value of another with an error having a standard deviation of about 5 per cent.

$$I_E = 1{\cdot}55\,\Phi - 1{\cdot}0 \times 10^{-3}\Phi^2 + 14 \tag{15}$$

$$I_{F2} = 1{\cdot}55 I_E - 1{\cdot}6 \times 10^{-3} I_E{}^2 - 156 \tag{16}$$

$$I_{F2} = 2{\cdot}34\,\Phi - 4{\cdot}5 \times 10^{-3}\,\Phi^2 - 142 \tag{17}$$

In equations (15) and (17), Φ is given in units of 10^{-22} W/m^2 per c/s.

It is interesting to note that if Φ is put equal to zero in equation (15), I_E would fall only to a value equivalent to about 12 per cent of its value at a low sunspot minimum. Assuming that the extrapolation to $\Phi = 0$ is valid and that Φ and I_E would be expected to fall to zero together, the residual value of I_E may represent a weak unidentified source of *E*-layer ionization other than the solar ionizing radiation represented by I_E.

Although the relation between I_E and Φ (equation 15) is slightly curved, the departure from linearity is not great enough to invalidate the discussion in Section 4.2 in which the relation between Ch_E and Φ was assumed to be linear.

5. Conclusions

Three indices of solar activity have been selected which appear to be free from seasonal variations; these are Φ, the flux of solar radio noise at $\lambda 10{\cdot}7$ cm, and I_E and I_{F2} which are based on the ionization in the *E*- and *F2*-layers of the ionosphere respectively (Appendices II, III and IV). I_E can be regarded as proportional to the flux of the ionizing radiation which is responsible for the *E*-layer, while I_{F2} is the component of the monthly mean sunspot number which is correlated with ionization in the *F2*-layer. The relation between I_{F2} and the intensity of the ionizing radiation is approximately linear.

Since both I_E and Φ represent solar fluxes of different kinds, it seems possible that they may have applications to the study of certain types of solar–terrestrial relations. In the preparation of medium-term forecasts of ionospheric behaviour, it would appear to be appropriate to use I_E or I_{F2}. The well-defined relations between the monthly mean values of I_{F2} and the *F2*-layer critical frequencies at all locations for which data are available provide an economical method of expressing f_{F2} as a function of location, season and solar activity on the lines of the preliminary investigation described by Minnis (1957).

Attempts to forecast the relative sunspot number R several months ahead have been made by many workers without great success. The main reason for this is the existence of a large and apparently random component in the month-to-month changes in R. Whether these fluctuations have any physical significance or not is doubtful in view of the arbitrary definition of R. The much more regular changes in I_E and I_{F2} make these indices less difficult to extrapolate.

It seems probable, however, that accurate forecasting of solar activity will not

be possible until more is known about how the various indices of solar activity are related to changes in the physical characteristics of the sun's surface and the visible features of the disk. Relations of this kind ought to be easier to discover using ionospheric and solar noise indices than those depending on the appearance and disappearance of sunspot groups.

Acknowledgements—The authors wish to acknowledge the considerable assistance given by Miss E. M. COTTENHAM in carrying out a large amount of numerical work and in the preparation of numerous graphs. The work was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

REFERENCES

ALLEN C. W.	1948	*Terr. Magn. Atmos. Elect.* **53,** 433.
APPLETON E. V. and NAISMITH R.	1939	*Phil. Mag.* **27,** 144.
APPLETON E. V. and PIGGOTT W. R.	1954	*J. Atmosph. Terr. Phys.* **5,** 141.
BEYNON W. J. G. and BROWN G. M.	1956	*Bull. Un. Radio Sci. Int.* **97,** 17.
GOODALL W. M.	1939	*Nature, Lond.* **143,** 977.
MEDD W. J. and COVINGTON A. E.	1958	*Proc. Inst. Radio Engrs., N.Y.* **46,** 112.
MINNIS C. M.	1955	*J. Atmosph. Terr. Phys.* **7,** 310.
MINNIS C. M.	1957	*Ann. Télécomm.* **12,** 164.
MINNIS C. M. and BAZZARD G. H.	1958	*Nature, Lond.* **181,** 1796.
PHILLIPS M. L.	1947	*Terr. Magn. Atmos. Elect.* **52,** 321.

VOL. 14 1959

APPENDIX I

A. *The* E*-layer character figure,* Ch_E

Values of Ch_E have been calculated for each month since January 1938 using the monthly mean values of f_E at noon at Slough (lat. 51°29′ N, long. 00°34′ W). f_E is read from the ionograms to the nearest 0·05 Mc/s but the monthly mean values have been given to the nearest 0·01 Mc/s. The values of sec χ used in equation (4) are the monthly mean values calculated using the data given by BEYNON and BROWN (1956).

B. *The E-layer index,* I_E

The relation between Ch_E and Φ for each month may be represented approximately by one of a series of linear expressions of the form

$$Ch_E = g_m + h_m\Phi \tag{A1}$$

If all months are considered together and seasonal variations in g_m and h_m are ignored, a single approximate expression may be used viz.

$$Ch_E = g_0 + h_0\Phi \tag{A2}$$

It follows from equation (A1) that, even if solar activity as indicated by Φ remains constant, Ch_E will vary in a manner depending on the seasonal changes in g_m and h_m.

The seasonal variations in Ch_E can be eliminated by correcting the values for each month in such a way that equation (A2) is valid for all months. The corrected values of Ch_E constitute a new index, I_E, which is defined by equations (A3) and (A4) and tabulated in Appendix II:

$$I_E = \lambda(m, \Phi)Ch_E \tag{A3}$$

$$\lambda(m, \Phi) = (g_0 + h_0\Phi)/(g_m + h_m\Phi) \tag{A4}$$

The tabulated values of I_E in Appendix II have been calculated by inserting in equation (A3) the values of Φ defined by equation (A1) and the measured values of Ch_E. In principle, equation (A3) could be expressed in terms of R instead of Φ but, since the correlation between Ch_E and Φ is much higher than between Ch_E and R, it is preferable to use Φ.

C. *The F2-layer index, I_{F2}*

The values of I_{F2} tabulated by MINNIS (1955) were derived using equations (7) and (8) and the monthly mean noon values of f_{F2} at Slough, Huancayo and Watheroo. This combination of locations was found to reduce to negligible proportions the geomagnetic control of I_{F2} referred to in Section 3.2.1.

Since the values given by MINNIS were calculated, the recent high levels of solar activity, particularly during northern winter, have provided valuable additional information on the relations between f_{F2} and R. It now appears that these relations do not depart significantly from linearity and the values of I_{F2} have been recalculated using linear relations and the most recent data which were available. The revised values are given in Appendix III.

APPENDIX II

The ionospheric index of solar activity, I_E

	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948
January	210	232	170	166	156	134	133	150	166	262	212
February	233	193	163	151	151	140	120	137	197	263	218
March	242	187	200	141	152	130	115	136	215	297	246
April	236	196	193	137	156	110	94	134	208	284	304
May	259	256	187	143	164	133	114	150	212	306	328
June	210	240	212	190	144	141	113	126	201	306	314
July	223	218	191	172	130	126	120	142	228	268	266
August	234	231	226	190	146	134	137	164	221	294	286
September	198	245	198	202	125	102	130	163	222	254	234
October	222	220	187	158	127	118	105	178	227	280	227
November	250	212	165	144	144	122	133	165	232	250	226
December	290	201	181	150	135	135	127	155	215	256	235

The ionospheric index of solar activity, I_E—continued

	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958
January	243	218	161	158	135	118	129	189	302	313
February	256	210	157	137	126	110	120	210	241	257
March	284	222	158	131	115	113	116	250	271	332
April	284	214	174	134	112	104	112	253	287	341
May	260	238	194	142	116	112	130	241	291	342
June	256	214	201	146	127	109	138	227	301	340
July	239	176	173	136	120	122	142	228	260	290
August	256	177	168	142	117	118	134	241	292	348
September	260	152	182	128	130	110	150	243	345	341
October	232	151	163	111	119	111	158	262	390	
November	252	161	165	133	116	117	174	298	342	
December	239	159	159	130	116	118	201	301	346	

Appendix III

The ionospheric index of solar activity, I_{F2}

	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947
January	102	128	66	61	40	16	1	19	40	125
February	113	91	68	48	41	14	1	28	71	146
March	124	87	78	46	49	26	11	24	91	149
April	120	94	84	35	57	31	−1	22	95	145
May	116	120	74	31	49	19	−8	29	97	168
June	120	118	78	43	25	18	−1	39	103	169
July	135	110	86	52	20	16	3	41	114	164
August	139	105	93	61	18	19	12	37	121	173
September	127	139	96	62	24	12	16	38	110	152
October	108	123	88	54	10	−1	21	59	115	141
November	119	105	78	48	24	−5	20	57	114	142
December	117	80	55	47	29	7	18	44	132	137

	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957
January	112	123	107	42	30	8	−1	4	71	143
February	107	145	106	42	29	10	4	9	88	125
March	114	148	102	45	25	−1	−5	4	112	140
April	(138)	148	112	72	32	5	1	3	122	144
May	(145)	138	116	72	18	6	−7	15	132	158
June	158	121	102	76	35	9	−12	28	119	161
July	146	118	85	78	35	10	−5	32	118	164
August	145	116	73	68	36	5	2	30	142	152
September	138	140	51	58	30	10	−3	37	146	158
October	129	129	57	57	22	9	7	50	155	(177)
November	114	132	50	42	16	7	10	73	(150)	(154)
December	116	121	37	47	13	0	−1	87	152	(156)

Figures in parentheses: ionospheric data available from two stations only.

VOL. 14 1959

Appendix IV

Monthly mean values of solar noise flux at λ10·7 cm (unit: 10^{-22} W . $m^{-2}(c/s)^{-1}$)

	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958
January		153	181	149	106	94	82	68	83	139	228	248
February	201	133	218	142	101	86	72	69	81	166	185	210
March	235	135	203	137	102	78	70	72	75	160	197	250
April	264	208	182	164	127	84	81	69	77	166	200	246
May	261	226	155	157	169	81	73	68	83	163	208	219
June	227	196	158	129	162	85	73	67	89	154	252	220
July	215	183	160	134	116	89	70	68	87	163	218	
August	231	173	175	121	110	93	75	70	91	194	202	
September	199	163	172	98	118	81	74	70	95	200	266	
October	208	158	177	99	105	82	72	73	111	200	281	
November	178	163	178	101	103	82	71	72	129	247	256	
December	174	190	163	99	101	84	70	74	132	249	282	

VOL.
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 229 to 235. Pergamon Press Ltd. Printed in Northern Ireland

"Chapman Behaviour" in the lower ionosphere

C. H. CUMMACK
Christchurch Geophysical Observatory, New Zealand

(*Received* 21 *July* 1958)

Abstract—It is shown that CHAPMAN's relations for the variation of critical frequency and height of a layer with solar zenith angle can be applied to any point on that layer and that it is not necessary to assume monochromatic radiation. These results are then applied to electron density vs. height profiles and it is shown that "Chapman behaviour" can be detected to a height of 160 km being consistent with a scale height of 10 or 15 km.

1. INTRODUCTION

CONSIDER a continuum of solar radiation acting on an isothermal atmosphere. The resulting electron production rate can be written down immediately using CHAPMAN'S (1931) well-known equations for monochromatic radiation:

$$q(z_0) = \int P(z) \exp \{1 - (z_0 - z) - Ch\,\psi \exp -(z_0 - z)\} \, . \, dz \qquad (1)$$

where

z_0 is the height (measured in units of scale-height) above some fixed datum level
$q(z_0)$ electron production rate at z_0
$P(z)$ an unknown weighting factor depending on the frequency distribution of the incident radiation and the radiation absorption coefficient
$Ch\,\psi$ is Chapman's function of solar zenith angle

In equation (1) each elementary layer, due to a single band of radiation, is referred to conditions at the sub-solar point, and, by displacing the origin by $\log Ch\,\psi$, may be referred to the maximum at the zenith angle under consideration. Equation (1) then becomes:

$$q(z_0 + \log Ch\,\psi) = \frac{1}{Ch\,\psi} \int P(z) \exp \{1 - (z_0 - z) - \exp (z_0 - z)\} \, . \, dz \qquad (2)$$

$$\equiv \frac{1}{Ch\,\psi} q_0(z_0) \quad \text{definition of} \quad q_0(z_0) \qquad (3)$$

It will be noticed that the integral in equation (2) is now independent of $Ch\,\psi$ and is constant provided $P(z)$ does not vary. If we introduce the equation of continuity:

$$\frac{dN}{dt} = q - \alpha N^2 \qquad (4)$$

where N is electron density
α recombination coefficient

then under quasi-equilibrium conditions:

$$\frac{dN}{dt} \ll q \quad \text{or} \quad \alpha N^2$$

$$N \doteq \left(\frac{q}{\alpha}\right)^{\frac{1}{2}} \tag{5}$$

and introducing equation (5) in equation (3) we get:

$$N(z_0 + \log Ch\,\psi) = Ch^{-\frac{1}{2}}\,\psi\,N_0(z_0) \tag{6}$$

provided α is approximately constant with height.

We see then that a layer due to non-monochromatic radiation formed in an isothermal atmosphere with large but uniform recombination coefficient preserves its shape and identity after being scaled by $Ch^{-\frac{1}{2}}\,\psi$ and lowered by $\log Ch\,\psi$, and also the well-known Chapman relations

$$N = N_0\,Ch^{-\frac{1}{2}}\,\psi \tag{7}$$

and

$$h = h_m + H \log Ch\,\psi \tag{8}$$

are not necessarily consequences of monochromatic radiation or refer only to the maximum of ionization.

As the experimental data is in the form of $h'(f)$ curves and $N\alpha f^2$ equation (6) may be rewritten:

$$f(h_0 + H \log Ch\,\psi) = Ch^{-\frac{1}{4}}\,\psi\,f_0(h_0) \tag{9}$$

where H is the scale height of the atmosphere and h_0 is the height in kilometres.

2. "Chapman" Behaviour in the Lower Ionosphere for Quasi-equilibrium Conditions

We now have a method for testing Chapman behaviour in the ionosphere. It is first necessary to know the scale height. Various workers (Mitra, 1952) have estimated a height of the order of 10 km for the E-layer. Here separate values of 5, 10, 15 and 20 km have been used.

Curves of electron density vs. height at Christchurch, New Zealand were reduced to true height curves by King's method (1954, 1956) for midday on one calm day each month of 1954; together with midday curves for each day of June 1954. Values of $f_0(h_0)$ for different heights and different scale heights were calculated using equation (9). The mean and variance for each group was computed and Chapman behaviour judged by the fit of the two curves. These results are shown in Fig. 1.

The curves show that the values $H = 10$ km and $H = 15$ km give the best fit while Chapman behaviour has fallen off above $h_0 = 160$ km.

3. Daily Variations

Previous workers (Mitra, 1952, p. 295) have found a good fit of E- and $F1$-layer critical frequency to the law $f = f_0\,Ch^{-\frac{1}{4}}\,\psi$ extending over large ranges of ψ. Hence it was decided to test the daily variations of the lower ionosphere by using equation

VOL
14
195

(9). June 1954 records were again chosen. All suitable records at 0800, 1000, 1200, 1400 and 1600 hours MMT were reduced and values of f for various heights computed. A scale height of 10 km was used. These results are plotted in Fig. 2.

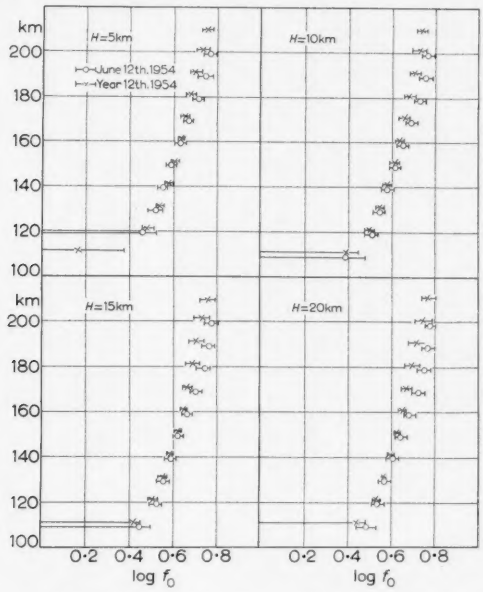


Fig. 1. Values of $f_0(h_0)$ for two selected groups of data during 1954.

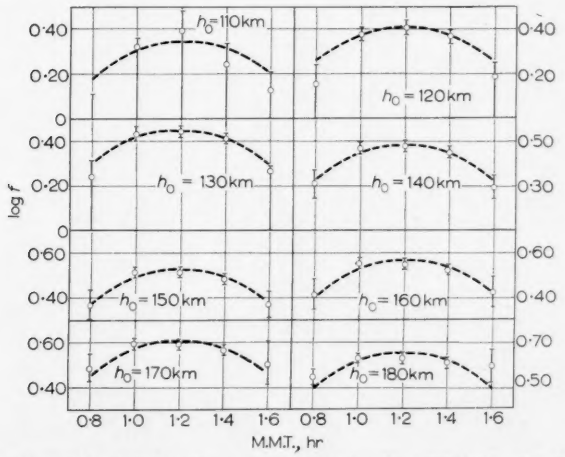
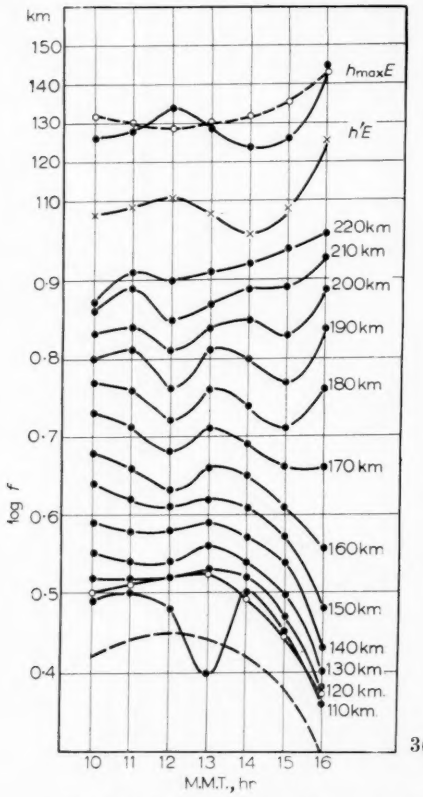


Fig. 2. Values of log f at a height of $h_0 + H \log Ch\ \psi$ for June 1954 showing daily variation. The dotted line give the expected form of the variation.

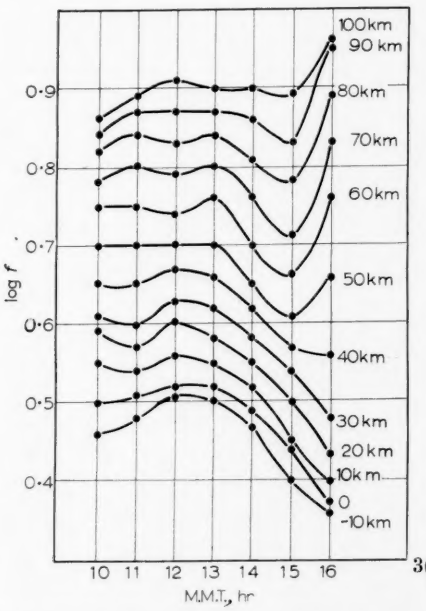
Here the fit is not so good and reasonable agreement is found only below 150 km. The poor agreement and large spread at 110 km may be attributed to the large gradient of electron density with height existing at the bottom of the E-layer. Any small error in height or slight perturbation will then make a large error in log f. The

VOL. 14 1959

general agreement with Chapman behaviour however is not as good as would be expected on considering the good fit of f_0E1 and f_0F1 to the law $f = f_0\, Ch^{-\frac{1}{4}}\, \psi$ as found by other workers. This may be due either to some physical process, such as ionospheric winds, effecting the general shape of the layers present but not greatly



3(a).



3(b).

Fig. 3(a) Daily variation of log f at a height of $h_0 + H \log Ch\, \psi$ for 1 May 1956. The lower dotted curve gives the expected form of variation for Chapman behaviour while the upper curves give the actual and expected height variations.

Fig. 3(b). The values of Fig. 3(a) replotted with $h_{\max} E1$ as basic height measure (f_0E1 is designated by 0 km.)

effecting their critical frequencies, or due to the inherent difficulties in computing true heights. Let us try and eliminate the second possible cause.

A record suitable for reduction to a true height curve should be free from sporadic-E and have good clean unambiguous traces free from oblique echoes. The ionograms of June 1954 were not entirely suitable for reduction but were as good as could be expected from a long series of records. The most likely cause of error in a reduction is an error in the minimum virtual height and such an error will be to a first approximation constant for all frequencies. $h'E$ has been read to a nominal accuracy of 1 km though the actual accuracy is likely to be of the order of 5 km. This error is large compared to the variation of H log $Ch\ \psi$ and could thus cause spurious variations of height.

As a check of the results given in Fig. 2 the records for the 1 May 1956 were examined. These were a particularly good set and had the following characteristics: (a) they were free of sporadic-E; (b) good clean unambiguous traces free from oblique echoes. (c) non-infinite $E1$ critical frequencies thus eliminating any possible effects due to valleys between layers.

In Fig. 3(a) values of $\log f$ at a height determined by equation (9) are plotted. Also plotted are values of $\log f_0 E1$ and $\log f_0\, Ch^{-\frac{1}{2}}\, \psi$ for an arbitrary value of f_0 while plotted in the upper curve are $h_{\max}\ E1$; $h'E1$ and $h_0 + \log Ch\ \psi$ for an arbitrary value of h_0. The same data are again plotted in Fig. 3(b) but this time each record is referred to $h_{\max}\ E1$ as the basic height measure thus eliminating their reliance on $h'E1$.

Fig. 3(a) shows considerable variation from that expected and it is impossible to give a range of height or ψ where Chapman behaviour is present. Furthermore as shown in Fig. 3(b) the differences are not explainable in terms of the most likely source of error (viz. $h'E1$). It then appears that differences are most likely explainable in terms of a perturbation of a Chapman layer.

4. Application to Early Sunrise-records

VOL. 14 1959

At Christchurch before ground sunrise the E-layer appears first as a ledge on the bottom of the F-layer which gradually descends and forms a separate layer, this one we will call the $E2$-layer. Close to ground sunrise the layer stratifies and the lower critical frequency increases until it usually blankets the $E2$-layer. This layer preserves its identity during the rest of the day as the $E1$ layer. Initially it was hoped to use this early morning $E2$-layer as a clue to the process of electron removal above 160 km but the results were paradoxical and seem worthwhile reporting. Using equation (5) and previous estimates of α, then q as a function of height can be computed. Midday curves and Rydbeck's (1946) values were used. By inserting these in equation (3) and using a scale height of 10 km the electron production rate at given heights near sunrise were found. The total number of electrons produced at a given height over a period was then compared with the number observed at the end of that period.

In Fig. 4 $\int^{\psi_0} q(\psi, h)\, dt$ integrated numerically for various values of h are compared with the N,h curves for $\psi = \psi_0$. Values of $q_0(z_0)$ were computed from the nearest suitable noon $h'f$ curves in order to minimize possible variations of $P(z)$.

The observed values of N are about an order of magnitude higher than that expected by extrapolation from the midday curves and show considerable variation between different mornings. A different value of scale height would not improve matters and it does not seem likely that the accepted value of recombination coefficient is out by a factor of ten. As the regular $E1$-layer forms as a stratification below this $E2$-region it seems unlikely that the extra ionization is produced by radiation normally absorbed in the D-layer at small values of ψ. A layer with a critical frequency smaller than 1 Mc (i.e. lower than the limit of the recorder) could affect the N,h curves but could not materially affect the general conclusions that the electron production rates are too high. It is possible that such a lower layer is present in the N,h curve for 2 September 1955.

5. TOTAL SOLAR FLUX

As was shown in the preceding section the electron production rate for conditions of overhead sun can be computed. We may use values of $q_0(h_0)$ below 160 km with some reliance and find the total electron production rate below this level. On the assumption that no appreciable number of secondary electrons are formed in the ionization process we may equate this to the number of solar photons absorbed in this region. Using midday June 1954 curves for $H = 10$ km $\int q_0(h_0) \cdot dh_0$ is equal to $1{\cdot}3 \times 10^9$ photons.

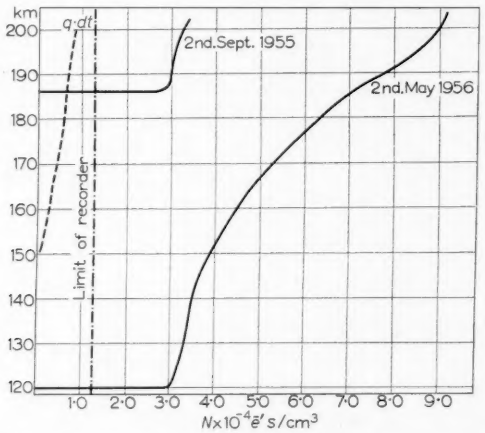


Fig. 4. Comparison of total number of electrons produced over a given sunrise period (as extrapolated from the noon curves) to the number actually observed at the end of that period.

By extrapolating the $q_0(h_0)$ curve to 180 km (this should not be seriously in error) and by using a mean scale height of 45 km above 180 km together with Bradbury's hypothesis as demonstrated by RATCLIFFE (1956) we may estimate the total ionizing photon flux for the ionosphere as $5{\cdot}7 \times 10^9$ photons. On assuming a 6000°K black-body sun the photon flux is $1{\cdot}0 \times 10^{10}$ photons/cm² per sec for energies above the first ionization potential of molecular oxygen and $9{\cdot}3 \times 10^8$ photons/cm² per sec for energies above the first ionization potential of atomic oxygen.

6. CONCLUSION

It has been found that it is possible to detect Chapman behaviour in the lower ionosphere from curves below 160 km being consistent with a scale height of 10 or 15 km. The agreement was not as close as that found with the *E*-layer critical frequency. The accuracy of the method however does not preclude Chapman behaviour of the *F1*-region which has its maximum above 160 km. It was found that departures from Chapman behaviour did exist which were more likely to be real perturbations than due to the inadequacy of height resolution of the equipment.

It has also been found that the pre-sunrise *E*-layer could not be explained in terms of Chapman behaviour in an isothermal atmosphere with constant recombination coefficient. No explanation is offered for this phenomenon.

From the derived N,h curves the total solar flux absorbed in the ionosphere

VOL
14
195

has been estimated. This indicates that the flux is not seriously different from that expected from a black-body at 6000°K.

Acknowledgements—This work is part of the research programme of the Geophysical Observatory of the New Zealand Department of Scientific and Industrial Research. The N,h curves were computed by Mr. G. A. M. KING, Mrs. R. MASON, Miss J. BRAND and myself. This paper was published by kind permission of the Superintendent, Mr. J. W. BEAGLEY.

REFERENCES

CHAPMAN S.	1931	*Proc. Phys. Soc. Lond.* **43,** 26.
KING G. A. M.	1954	*J. Atmosph. Terr. Phys.* **5,** 245.
KING G. A. M.	1956	*J. Atmosph. Terr. Phys.* **8,** 184.
MITRA S. K.	1952	*The Upper Atmosphere.* The Asiatic Society Monograph Series, Vol. V.
RATCLIFFE J. A.	1956	*Phil. Trans.* **248,** 621.
RYDBECK O.	1946	*Chalmers Tek. Högsk. Handl.* No. 53.

VOL.
14
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 236 to 243. Pergamon Press Ltd. Printed in Northern Ireland

Some deductions of ionospheric information from the observations of emissions from satellite 1957α2—I

The theory of the analysis

G. J. AITCHISON and K. WEEKES
Cavendish Laboratory, Cambridge

(*Received* 31 *July* 1958)

1. INTRODUCTION

THE launching of the first artificial earth satellite (1957α) by the USSR provided an opportunity of obtaining information about the ionosphere at heights above the *F2*-peak. The satellite emitted radio waves of frequency 20 and 40 Mc/s which were appreciably modified during their transmission through the ionosphere to the ground. It should therefore be possible to obtain information about the distribution of ionization by making suitable observations on the received wave. It turns out that the most useful quantities for this purpose are the Doppler frequency shift and the fading of the signal caused by the magneto-optical Faraday effect. It is the purpose of this paper to develop an approximate theory of these two phenomena and to use it to interpret some observations.

VOL
14
195

1.2. *The Faraday effect*

The geomagnetic field makes the ionosphere doubly refracting so that there are, at any level, two characteristic waves, generally elliptically polarized, with different refractive indices. Thus as a plane polarized wave is transmitted through the ionosphere its plane of polarization is rotated and by analogy with the corresponding phenomenon in optics this rotation is often called the "Faraday rotation". As the position of the source alters, the amount of rotation alters, and the signal received on a linearly polarized aerial exhibits fading at a rate determined by the rate of change of the phase difference between the two characteristic waves. Now the Doppler frequency shift of one of the waves is determined by the rate of change of its phase path so that the rate of Faraday fading is equal to the difference of the Doppler shifts for the two waves.

In describing the rate of Faraday fading it is important to distinguish between the number of complete rotations of a rotating linearly polarized wave and the number of fades on a linearly polarized aerial. One of these is twice the other and this difference accounts for an apparent difference of a factor of 2 between the expressions given by BROWNE *et al.* (1956) and EVANS (1956).

In Part I of this paper we develop the theory of the effect of the ionosphere on the Doppler frequency shift and on the frequency of Faraday fading. As it is impossible to measure an instantaneous frequency we have preferred to develop the theory in terms of the difference in phase path of the waves for two positions of the source.

Browne *et al.* (1956) have derived a simple expression for the Faraday fading by neglecting refraction of the waves and calculating the difference in phase path of the two characteristic waves along the straight line from source to receiver. It is not obvious how far this approximation is valid, particularly for the lower frequency waves emitted by the satellite, and it is part of the object of this paper to determine the range of validity of the approximations.

In Part II (Aitchison, Thomson and Weekes, 1958) the approximate expressions developed in Part I are applied to some observations made on the emissions from satellite 1957α2.

2. The Phase Path

2.1. *Calculation of the phase path*

The evaluation of the effect of the ionosphere on the Doppler shift and on the Faraday fading involves only the calculation of the phase path along the ray from the satellite to the observer. The accurate calculation of the path of the ray through the inhomogeneous ionosphere, rendered birefringent by the earth's magnetic field, is difficult, in particular it is not possible to apply a simple Snell's law. For satellite 1957α in its transits over England the emitted frequencies were high compared to the plasma frequency at the *F2*-peak and compared to the electron-gyrofrequencies, and under these conditions it is possible to obtain a simple approximate solution of the problem.

Appleton's magneto-ionic theory leads to the general expression for the refractive index, in the usual notation:

$$(\mu - ic\kappa/\omega)^2 = 1 - \frac{X}{1 - iZ - \dfrac{\frac{1}{2}Y_T^2}{1 - X - iZ} \pm \sqrt{\left(\dfrac{\frac{1}{4}Y_T^4}{(1 - X - iZ)^2}\right) + Y_L^2}}$$

For satellite 1957α on the occasions with which we are concerned the value of X on 20 Mc/s was about 0·1, $Z \sim 10^{-5}$ and $Y \sim 0{\cdot}05$.

In these conditions

$$Y_L^2 \gg Y_T^4/4(1 - X - iZ)^2$$

provided

$$\cos\phi > \tfrac{1}{2}Y \sin^2\phi$$

or

$$\phi < 80^\circ$$

where ϕ is the angle between the ray and the magnetic field. Thus, for all conditions in these experiments, the quasi-longitudinal approximation is valid and

$$(\mu - ic\kappa/\omega)^2 = 1 - X/(1 \pm Y_L - iZ)$$

And since Z is very small:

$$\mu^2 = 1 - \omega_0^2/\omega(\omega \pm \omega_H \cos\phi) \qquad (1)$$

Since $\omega_H/\omega \sim 0{\cdot}05$ it is a close approximation to assume that μ is independent of ϕ and to apply Snell's law to trace the ray.

If we suppose that the ionosphere is horizontally stratified as in Fig. 1, the phase path P is given by

$$P = \int_{r_0}^{r_s} \mu \, ds = \int_{r_o}^{r_s} \mu \sec\psi \, dr$$

From Snell's law for a spherical stratified medium

$$\mu r \sin \psi = r_0 \sin i$$

Hence

$$P = \int_{r_0}^{r_s} \mu^2 r (\mu^2 r^2 - r_0^2 \sin^2 i)^{-1/2} \, dr \tag{2}$$

Also the co-ordinate θ_s of the satellite is related to the angle of incidence, i, by the expression:

$$\theta_s = \int_{r_0}^{r_s} (r_0 \sin i / r)(\mu^2 r^2 - r_0^2 \sin^2 i)^{-1/2} \, dr \tag{3}$$

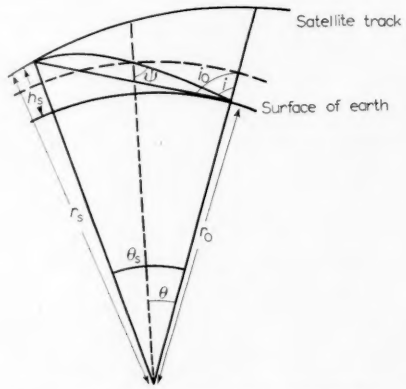


Fig. 1. Diagrammatic representation of the satellite track and the notation used.

Equations (2) and (3) then determine in principle the values of i and P for any model of the ionosphere.

2.2. *The change of phase path with μ: first approximation*

The difference between the Doppler shifts on 20 Mc/s and 40 Mc/s and the Faraday fading on a single frequency depend on the difference ΔP in P for slightly different values of μ. We may conveniently calculate this difference from the expressions in equations (1) and (2). We may write:

$$\Delta P = \int_{r_0}^{r_s} \left\{ \frac{\partial}{\partial \mu} \left(\frac{\partial P}{\partial r} \right) \delta\mu + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \left(\frac{\partial P}{\partial r} \right) \delta\mu^2 + \ldots \right\} dr \tag{4}$$

$$\Delta\theta_s = 0 = \int_{r_0}^{r_s} \left\{ \frac{\partial}{\partial \mu} \left(\frac{\partial \theta}{\partial r} \right) \delta\mu + \frac{1}{2} \frac{\partial^2}{\partial \mu^2} \left(\frac{\partial \theta}{\partial r} \right) \delta\mu^2 + \ldots \right\} dr \tag{5}$$

To a first approximation, the range of validity of which we will discuss later, we may neglect the second-order terms in these expressions and write:

$$\Delta P = \int_{r_0}^{r_s} \frac{\partial}{\partial \mu} \left(\frac{\partial P}{\partial r} \right) \delta\mu \, dr \tag{6}$$

$$\Delta\theta_s = 0 = \int_{r_0}^{r_s} \frac{\partial}{\partial \mu} \left(\frac{\partial \theta}{\partial r} \right) \delta\mu \, dr \tag{7}$$

VOL
14
195

From equations (2) and (3), differentiating with respect to μ, we obtain:

$$\Delta P = \int_{r_0}^{r_s} \mu r(\mu^2 r^2 - r_0^2 \sin^2 i)^{-1/2}\, \delta\mu\, dr + \int_{r_0}^{r_s} \frac{r_0^2 \sin^2 i \left(\cot i \frac{\partial i}{\partial \mu} - \frac{1}{\mu}\right)\delta\mu}{r(\mu^2 r^2 - r_0^2 \sin^2 i)^{3/2}}\, dr$$

$$\Delta\theta_s = 0 = \int_{r_0}^{r_s} \frac{r_0 \sin i \left(\cot i \frac{\partial i}{\partial \mu} - \frac{1}{\mu}\right)\delta\mu}{r(\mu^2 r^2 - r_0^2 \sin^2 i)^{3/2}}\, dr$$

Hence

$$\Delta P = \int_{r_0}^{r_s} \mu r(\mu^2 r^2 - r_0^2 \sin^2 i)^{-1/2}\, \delta\mu\, dr + r_0 \sin i\, \Delta\theta_s$$

$$= \int_{r_0}^{r_s} (1 - r_0^2 \sin^2 i/\mu^2 r^2)^{-1/2}\, \delta\mu\, dr \qquad (8)$$

3. Application to the Doppler and Faraday Frequencies

The difference between the phase paths on frequencies $2f$ and f is given by equation (8) with $\delta\mu$ given by equation (1) in which we may neglect the term $\omega_H \cos\phi$ to our approximation, then

$$\delta\mu = \frac{\omega_0^2}{8\pi^2}\left(\frac{1}{f^2} - \frac{1}{4f^2}\right)$$

$$= \frac{3}{8}\frac{f_0^2}{f^2}$$

The Doppler shift Δf in frequency f is given by

$$\Delta f = -\frac{1}{\lambda_0}\frac{\partial}{\partial t} P$$

or

$$c\frac{\Delta f}{f} = -\frac{\partial}{\partial t} P$$

Thus the difference in the Doppler shifts on frequencies f and $2f$ is related to the differences in the phase path by the relation

$$2\Delta f - \Delta(2f) = \frac{2f}{c}\frac{\partial}{\partial t}\Delta P$$

$$= \frac{3}{4}\frac{f}{c}\frac{\partial}{\partial t}\int_{r_0}^{r_s} \frac{f_0^2}{f^2}\left(1 - \frac{r_0^2 \sin^2 i}{\mu^2 r^2}\right)^{-1/2} dr \qquad (9)$$

For the Faraday fading frequency we are concerned with the difference in P for the two magneto-ionic components and from equation (1)

$$\delta\mu = \frac{\omega_0^2}{2\omega}\left(\frac{1}{\omega - \omega_H \cos\phi} - \frac{1}{\omega + \omega_H \cos\phi}\right)$$

$$\doteq \frac{\omega_0^2}{\omega^3}\,\omega_H \cos\phi$$

Now

$$\omega_0^2 = \frac{4\pi N e^2}{\varepsilon_0 m} \qquad \text{and} \qquad \omega_H = \frac{eB}{m}$$

The Faraday fading frequency

$$F = \frac{1}{\lambda}\frac{\partial}{\partial t}\Delta P$$

Hence

$$F = \frac{e^3}{2\pi^2\,\varepsilon_0 m^2 c f^2}\frac{\partial}{\partial t}\int_{r_0}^{r_s} BN\cos\phi\,\{1 - r_0^2\sin^2 i/\omega r^2\}^{-1/2}\,dr \tag{10}$$

and

$$2\Delta f - \Delta(2f) = \frac{3e^2}{4\pi\,\varepsilon_0 mcf}\frac{\partial}{\partial t}\int_{r_0}^{r_s} N\{1 - r_0^2\sin^2 i/\mu\omega r^2\}^{-1/2}\,dr \tag{11}$$

The important ionization at night, the time in which we are interested, lies between some 200 km and 500 km (the height of the satellite) above the earth and in this region B may be assumed constant and as μ is not very different from unity, ϕ is approximately independent of r, consequently the two integrals in equations (10) and (11) are essentially the same.

4. Approximations to the Integrals

To evaluate the frequency of Faraday fading or the difference in the Doppler frequencies for f and $2f$, it is in general necessary to evaluate the integral in equation (11) and that in equation (3) which gives the value in $\sin i$. This involves the assumption of a model of the ionosphere which is then checked by the predicted fading frequencies. However, for the cases of interest to us it is possible to make an approximate evaluation of the integrals for small values of $\sin i$.

If i_0 be the angle of incidence corresponding to a straight line from the receiver to the satellite, we may write:

$$\sin i = \sin i_0(1 - \delta)$$
$$r = r_0(1 + h/r_0)$$

Then the integral in equation (11) becomes:

$$\int_0^{h_s} N\left\{1 - \sin^2 i_0(1 - 2\delta)\Big/\left(1 - \frac{\omega_0^2}{\omega^2}\right)\left(1 + \frac{2l}{r_0}\right)\right\}^{-1/2} dh$$

$$= \int_0^{h_s} N\sec i_0\left\{1 + \tan^2 i_0\left(2\delta - \frac{\omega_0^2}{\omega^2} + \frac{2h}{r_0}\right)\right\}^{-1/2} dh$$

$$\fallingdotseq \int_0^{h_s} N\sec i_0\,dh - \int_0^{h_s} N\sec i_0\tan^2 i_0\left(\delta + \frac{h}{r_0} - \frac{\omega_0^2}{2\omega^2}\right) dh$$

Thus to a first approximation, valid for sufficiently small i_0, we may write equations (10) and (11) in the form

$$F = \frac{e^3 B}{2\pi^2\,\varepsilon_0\,mcf^2}\frac{\partial}{\partial t}\left\{\cos\phi\sec i_0\int_0^{h_s} N\,dh\right\}$$

$$2\Delta f - \Delta(2f) = \frac{3e^2}{4\pi\,\varepsilon_0 mcf}\frac{\partial}{\partial t}\left\{\sec i_0\int_0^{h_s} N\,dh\right\}$$

For frequencies of 20 Mc/s and 40 Mc/s these expressions lead to the numerical values:

$$F_{20} = 7{\cdot}32 \,.\, 10^{-12} \frac{\partial}{\partial t} \left\{ \cos \phi \sec i_0 \int_0^{h_s} N \, dh \right\} \tag{13}$$

$$2\Delta f_{20} - \Delta f_{40} = 1{\cdot}01 \,.\, 10^{-10} \frac{\partial}{\partial t} \left\{ \sec i_0 \int_0^{h_s} N \, dh \right\} \tag{14}$$

An expression equivalent to (13) has been derived by other workers (Browne *et al.*, 1956) by neglecting any refraction and supposing that the ordinary and extraordinary rays follow the same (straight line) path and that the difference in phase path is given by this path multiplied by the difference in the refractive indices.

5. The Validity of the Approximations

In deriving equations (13) and (14) we have made several approximations and it is important to estimate the errors involved in these approximations and the conditions in which the simple equations are valid for useful deductions of the ionization.

(1) It is fundamental to the whole argument that we have assumed that μ is independent of direction—for the frequencies in use ω_H/ω is about 0·05, but if ω_0/ω is not near unity the bending is in any case small and the dependence on direction is unlikely to cause any serious error.

(2) We have assumed spherical stratification of the ionization. For small angles of incidence so that the path of the ray is always within about 500 km of the receiving point this may be a fair approximation, but when the satellite covers great distances while under observation it is probable that horizontal gradients of ionization may be important.

To remove these two assumptions would involve a detailed numerical calculation of ray tracing in a model ionosphere which could be assumed to vary horizontally in accordance with vertical incidence sounding data for the actual time of the passage of the satellite.

(3) In deriving equation (11) we have neglected $\omega_H \cos \phi$ in comparison with ω. In the cases considered this is an error of the order 5 per cent. However it may be noted that if the satellite is designed to emit circularly polarized waves the polarization of the wave received is uncertain and the correct value for μ is likewise uncertain. Thus rotation of the satellite might convert the received wave from the ordinary component to the extraordinary with corresponding distortion of the Doppler curve.

For the transits studied here, the Faraday effect produced about 40 fading c/min on the radiofrequency of 20 Mc/s so that the difference in the Doppler frequency shifts for the ordinary and extraordinary waves was about $\frac{2}{3}$ c/s. The change in phase path in 30 sec was used to deduce ionospheric information and this differed from the geometrical change in the path by about 250 wavelengths, uncertainty in the polarization would therefore cause an error of about 10 per cent. However it has been inferred from the observations that the satellite was rotating at 13 rev/min and hence completes 6·5 rev in the 30 sec interval, thus the total

VOL. 14 1959

change in phase path would be expected to be that appropriate to the average refractive index:

$$\mu_{\text{av}} = 1 - \frac{\omega_0^2}{2\omega^2(1 - \omega_H^2 \cos^2 \phi/\omega^2)}$$

in which the term involving ω_H contributes an error of only 0·25 per cent.

(4) Neglect of the second terms in equations (4) and (5) for the Taylor expansions of ΔP and $\Delta\theta$ may cause appreciable error particularly at large angles of incidence. It is difficult to estimate the magnitude of the error for the general case, but if the ionosphere is considered as a uniform slab of ionization between the satellite and some lower level, then it may be shown that the ratio of the second term in equation (4) to the first is of order

$$\int_{r_0}^{r_s} -\delta\mu/2(1 - r_0^2 \sin^2 i/\mu^2 r^2)^{1/2}\, dr$$

If the denominator is evaluated for the maximum of the *F2*-layer, then

$$\mu^2 \sim 1 - \omega_0^2/\omega^2 \qquad r^2 = r_0^2\left(1 + \frac{2h_m}{r_0}\right)$$

and for cases in the early morning $\omega_0/\omega \sim 1/3$, $h_m \sim 400$ km then

$$\omega^2 r^2 \simeq 1{\cdot}01 r_0^2$$

Hence the proportionate error is roughly $\delta\omega/2 \cos^2 i$. For the Doppler case (20 Mc/s and 40 Mc/s) $\delta\mu \sim 1/24$ and the proportionate error is 0·04 at 45° and 0·08 at 60° angle of incidence.

For the Faraday frequency $\delta\mu \sim \omega_0^2\omega_H/\omega^3 \sim 1/180$ and the error is 5 per cent at an angle of incidence of 72°.

(5) In evaluating the integral in equations (10) and (11) the second and higher terms in equation (12) have been neglected. This involves neglecting

$$\tan^2 i_0(\delta + h/r_0 - \omega_0^2/2\omega^2)$$

in comparison with unity. For the slab model of the ionosphere it may be shown that $\delta \sim \omega_0^2/\omega^2$ so that we get an overestimate of the error by considering the magnitude of $h \tan^2 i_0/r_0$, at the satellite $h/r_0 \sim 1/13$ and the average value for the ionospheric *F2*-region might be taken to be $\sim$1/16. Thus the error involved in neglecting this term is about 6 per cent at an angle of incidence of 45°.

Thus the simple formula (13) for the Faraday fading rate will be correct to 6 per cent up to angles of incidence of about 45°; equation (14) for the Doppler effect is liable to be in error by 10 per cent at this angle of incidence. The error is such that the value for $\int N\, dh$ deduced at the larger angles of incidence will be an underestimate.

6.

It is not practicable to measure either the Doppler frequency shift or the Faraday fading frequency instantaneously, what is in fact done is to determine the number of cycles of the oscillation between two times and thus determine the average frequency. In these circumstances it is most convenient to integrate

equations (13) and (14) with respect to time and in fact work with the differences in phase path, ΔP, for the two frequencies used in the Doppler method (equation 14) and for the two magneto–ionic components in the Faraday fading (equation 13).

If we denote by $[n_D]_1^2$, $[n_F]_1^2$, the total number of cycles of the Doppler shift frequency and the Faraday fading when the satellite moves from position 1 to position 2, we may rewrite equations (13) and (14) in the form:

$$[n_{F_{20}}]_1^2 = 7{\cdot}32 \,.\, 10^{-12}\left\{(\cos\phi_2 \sec i_2 - \cos\phi_1 \sec i)\int_0^{h_1} N\,dh + \right.$$
$$\left. + \cos\phi_2 \sec i_2 \int_{h_1}^{h_2} N\,dh\right\} \quad (15)$$

$$\left[2n_{D_{20}} - n_{D_{40}}\right]_1^2 = 1{\cdot}01 \,.\, 10^{-10}\left\{(\sec i_2 - \sec i_1)\int_0^{h_1} N\,dh + \sec i_2 \int_{h_1}^{h_2} N\,dh\right\} \quad (16)$$

Now in these latitudes the satellite was not even approximately travelling horizontally so that $h_1 \neq h_2$ and it is not possible to determine $\int_0^h N\,dh$ from either (15) or (16) alone, but clearly if both equations could be used it would be possible to determine separately $\int_0^{h_1} N\,dh$ and $\int_{h_1}^{h_2} N\,dh$.

VOL. 14 1959

To measure the Doppler shift it is necessary to know the frequency of the sender in the satellite; because the height is changing the Doppler shift is not zero at the distance of closest approach but at some place, close to that position, at which the ray from the receiver is perpendicular to the satellite track; it is very difficult to determine this point, moreover the frequency of the sender may alter during the transit. It would be possible, in principle, to eliminate the difficulties by use of simultaneous Doppler measurements at two places.

The accuracy of the Doppler records for satellite 1957α1 was not adequate to allow a detailed study of the effects to be made. But automatic Doppler records on future satellites should make it possible to use these equations to deduce values of $\int N\,dh$ and hence the distribution of N with height at the levels of the satellite.

References

Browne I. C., Evans J. V., Hargreaves J. K. and Murray W. A. S. 1956 *Proc. Phys. Soc. Lond.* B **69**, 901.

Evans J. V. 1956 *Proc. Phys. Soc. Lond.* B **69**, 953.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 244 to 248. Pergamon Press Ltd. Printed in Northern Ireland

Some deductions of ionospheric information from the observations of emissions from satellite 1957α2—II

Experimental procedure and results

G. J. Aitchison, J. H. Thomson and K. Weekes
Cavendish Laboratory, Cambridge

(*Received* 31 *July* 1958)

1. The Available Data

For the application of the approximate analysis described in Part I, it was necessary for the satellite track to pass sufficiently near the receiving site for the angle of incidence to be less than about 45° for a period of about 30 sec. The tracks which satisfied these conditions fell naturally into two groups. In one the height was 200–220 km, well below h_mF2, and since it occurred at night the ionization in E-region was very small, and in the other the height was about 500 km, well above h_mF2, and the time of observation was in the early morning.

The high transits for which the satellite passed most nearly overhead were 0531 hours 11 October, 0531 hours 12 October, 0339 hours 18 October and 0335 hours 19 October. The assumption of horizontal stratification of the ionosphere is doubtful for the 0530 group as this is rather near sunrise and consequently the analysis was limited to the 0330 group of transits for 17–22 October. The number of results available was so small that some records were analysed for intervals for which the angle of incidence appreciably exceeded 45°.

2. Analysis of the Faraday Fading Records

The senders were not keyed during the period concerned and from the records of the amplitude of the signals received on both frequencies taken at Cambridge it was easily recognized that fading arose from any or all of three causes.

(a) Movement of the satellite through, or past, irregularities such as are believed to produce "scintillation" of radio stars. This aspect of the fading has been discussed by our colleague (Kent 1958). Fast fading of this type, when present, frequently masked the Faraday fading.

(b) Rotation of the satellite itself. The Faraday fading would be expected (see equation 10) to be four times as fast on a frequency of 20 Mc/s as on 40 Mc/s. During the evening transits, when the satellite was below nearly all the ionization in the ionosphere, fading occurred at a rate of 13·0/min on both the emitted frequencies. It was concluded that this fading was not ionospheric in origin and that it was caused by the rotation of the satellite.

(c) The Faraday fading. During the high (morning) transits the rates of fading were different on the two frequencies. The procedure adopted for analysis of the records was to count the number of fading cycles in a period of 30 sec. This number was typically between 20 and 30 for the records on 20 Mc/s and the number could be estimated to roughly 0·1 c. The counts were made for both frequencies.

If the satellite were not rotating the ratio of the counts should be 4. It was consistently found that the ratio was less than 4 but that it approximated closely to the theoretical value if 6·5 c were subtracted from each to allow for the fading caused by the rotation. In a typical case for a 30 sec interval 24·9 c were counted on 20 Mc/s and 11·0 on 40 Mc/s, so that (24·9 − 6·5) = 18·4 and (11·0 − 6·5) = 4·5 represent the frequencies of the Faraday fading: they are seen to be closely in the ratio 4 : 1. Five of the six transits from which results have been obtained gave similar agreement when 6·5 cycles were subtracted from the 30 sec counts; in the other case (21 October) for which there was no 20 Mc/s record it was assumed that the same correction was needed.

Unpublished measurements made at Cambridge, of the time horizontal range and height of the satellite at the individual meridian transits and the orbital data published by the staff of the Mullard Radio Astronomy Observatory (1957) were used to calculate the values of i and φ at the beginning and end of each $\frac{1}{2}$ min interval. Direct substitution in equation (15) of Part I (Aitchison and Weekes, 1958) then gave the results in Table 1 which were reduced from the 20 Mc/s records for the two transits for 19 October and 20 October.

It is interesting to compare the values of the integrated electron content deduced from the Faraday fading with that calculated from ionograms recorded at the same time. The ionograms used were those recorded hourly at the Radio Research Station, Slough, and they were analysed to provide $N(h)$ curves which showed the electron density N as a function of height h up to the peak of the $F2$-layer. The values of h_mF2, N_mF2 and $\int_0^{h_mF2} Ndh$ at the times of transit were obtained by linear interpolation.

Table 1

Time	h_mF2 (km)	N_mF2 (cm^{-3})	I ($\times 10^{12}$ cm^{-2})	Magnitude of I from Faraday fading ($\times 10^{12}$)	Magnitude of I from model ionosphere ($\times 10^{12}$)
Oct. 19, 1957 0335 hours	390	$5{\cdot}8 \times 10^5$	$4{\cdot}1 \int_0^{500} Ndh \pm 7{\cdot}0 \int_{500}^{512} Ndh$	15·2	15·4
			$\int_0^{512} Ndh \pm 8{\cdot}0 \int_{512}^{524} Ndh$	14·2	16·8
			$\int_0^{524} Ndh \pm 8{\cdot}7 \int_{524}^{536} Ndh$	16·0	18·0
Oct. 20, 1957 0330 hours	400	$5{\cdot}6 \times 10^5$	$5{\cdot}0 \int_0^{518} Ndh \pm 7{\cdot}5 \int_{518}^{530} Ndh$	16·6	16·6
			$\int_0^{530} Ndh \pm 8{\cdot}3 \int_{530}^{542} Ndh$	16·6	17·9

For a first comparison with the satellite results it was supposed that the ionization density at heights between the *F2*-peak and the satellite had the value measured at the peak. The magnitude which would be assumed by the quantity I of column 5 of Table 1 if this model ionosphere were correct is shown in column 6.

The fact that the values of I deduced from the Faraday fading are slightly less than those deduced from the model ionosphere shows that the assumption that the electron density does not decrease above the peak is incorrect. A second model was therefore considered in which the electron density was supposed to decrease linearly above h_mF2, so that it had fallen to $(1 - \alpha)N_m$ at a height 100 km above the peak.

The values deduced from α are shown in Table 2; some additional results have been included from records in which only 40 Mc/s was observed. Some of the results apply to intervals for some part of which the angle of incidence increased 45°. These results are marked with an asterisk. All heights in the table are given to the nearest 5 km.

Table 2

Date (1957)	h_mF2	$\int_0^{h_mF2} Ndh$ (10^{12} cm^{-2})	N_mF2 ($\times 10^5$ cm^{-3})	Frequency (Mc/s)	Height of satellite (km)	Height above h_mF2	α
17 Oct.	355	2·9	4·7	40	475	120	0·00
18 Oct.	365	4·2	5·8	40	450	85	0·09*
				40	495	130	0·32
19 Oct.	390	4·1	5·8	40	475	85	0·00
				20	500	110	0·20
				20	510	120	0·22
				40	520	130	0·19
				20	525	135	0·14*
20 Oct.	400	5·0	5·6	40	475	75	0·00*
				20	520	120	0·00
				20	530	130	0·10
				40	550	150	0·12*
21 Oct.	425	3·3	4·2	40	505	80	0·00
22 Oct.	420	2·9	3·1	40	525	105	0·00

3. The Doppler Records

Estimates of the electron content may also be made using equation (16) Part I (Aitchison and Weekes) if simultaneous records are available for the Doppler frequency shift on both frequencies radiated from the satellite. In Fig. 1 a typical record of the received frequency is plotted for the two frequencies, with the frequency scale doubled for the lower curve. If there were no ionosphere these two curves would be identical and the ionospheric effect has to be evaluated as the difference between them. Thus $[n_{D_{40}}]_1^2$ is represented by the area PQR on the figure and $2[n_{D_{20}}]_1^2$ by the area XYZ. Under the conditions prevailing these areas are very nearly equal. In a typical $\frac{1}{2}$ min interval the areas correspond roughly to 5000 c, and it would be expected that, because of propagation through the

ionosphere there would be a difference of about 260 c between the two areas. To measure this difference to an accuracy of ± 2 per cent it would be necessary to estimate the two areas individually to about 1 in 2000.

The best simultaneous records available to us were those taken at the Post Office Radio Stations at Baldock and Banbury where accurate readings of frequency were taken at intervals of about 10 sec by manual observations.

For a number of the best of these records (where the points lay on a smooth

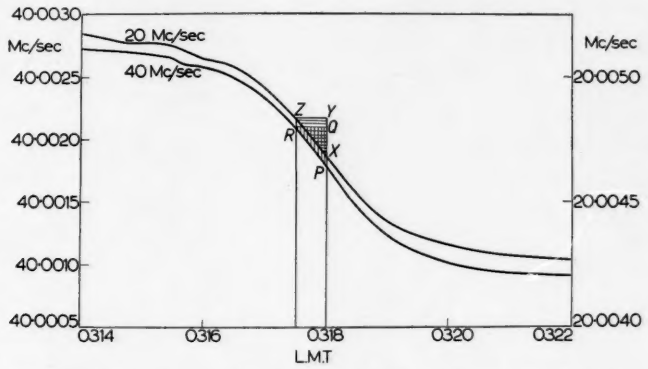


Fig. 1. Typical Doppler curves for 20 and 40 Mc/s signals. The effect of the ionosphere is proportional to the difference of the shaded areas.

curve) the differences of the areas were calculated for the $\frac{1}{2}$ min intervals from the time of steepest slope; these lay between 200 and 300 c. There is some error in assuming that the Doppler shift is zero at the time of most rapid change so that these values may be taken as consistent with the results from the Faraday fading within the accuracy of the Doppler results.

A more detailed study of the Doppler results has been made by HIBBERD and THOMAS (1958) by a method which does not impose the same restrictions on the angle of incidence as the approximate method used here.

4. CONCLUSIONS

(1) The analysis of the Faraday fading suggests that the ionization density above the maximum of *F2* at night decreases only slowly and that at 100 km above $h_m F2$ it has not decreased by more than 20 per cent. These results are tentative because of the inaccurate knowledge of the orbit for 1957α2 and because the analysis itself is only approximate.

(2) The best Doppler records, although not sufficiently accurate for detailed analysis, do not appear to be in disagreement with these conclusions.

(3) If full advantage is to be taken of the information available in the transmissions from future satellites it is highly desirable that accurate orbit data should be available as a result of visual observations and that records of the Doppler effect should be made continuously and automatically at several receiving sites.

Acknowledgements—The signal strength records on which the analysis in this paper was based were obtained by the combined efforts of members of the Mullard Radio Astronomy Observatory and of the ionosphere physics group of the Cavendish Laboratory. We are very much indebted to the Post Office Engineering Department for the provision of the Doppler frequency data on both frequencies and to the Director of Radio Research, DSIR, for supplying vertical incidence ionosphere records for analysis.

One of us (G. J. A.) is indebted to the University of Adelaide for financial assistance and leave of absence for a period of study at the Cavendish Laboratory.

References

Aitchison and Weekes 1958 *J. Atmosph. Terr. Phys.* This issue, p. 236.

Kent G. S. 1959 *J. Atmosph. Terr. Phys.* In press.

Mullard Radio Astronomy Laboratory Staff 1957 *Nature, Lond.* **180,** 879.

Thomas J. A. and Hibberd F. H. 1959 *J. Atmosph. Terr. Phys.* **13,** 376.

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 249 to 261. Pergamon Press Ltd. Printed in Northern Ireland

Coefficients for the rapid reduction of h'–f records to N–h profiles without computing aids*

E. R. SCHMERLING and C. A. VENTRICE
Ionosphere Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania

(*Received* 11 *August* 1958)

Abstract—Tables of coefficients are presented by means of which h'–f records may be readily reduced to electron-density–height profiles without the use of computing aids. The tables presented are for any station whose magnetic dip angle does not exceed 80°. The ordinary ray trace is utilized. No special assumptions concerning profile shapes are made. Account is taken of the earth's magnetic field, but collisions are neglected.

The sensitivity of these coefficients to magnetic dip angle and gyrofrequency is discussed. Sample h'–f records are reduced by means of the coefficients and the results are compared with those from the Budden matrix method.

1. INTRODUCTION

IN this paper tables of 5 and 10-point coefficients are presented, by means of which ionospheric h'–f records (ionograms) may be readily reduced to N–h (electron–density–height) profiles without further computational aids. Full account is taken of the earth's magnetic field, and no assumptions are made concerning layer shape, other than a monotonic variation of electron density with height. Electronic collisions are neglected, and a ray-theory approach (utilizing the ordinary ray) is used. The tables are consequently of value mainly for region F. These tables are valid for all observing stations where the magnetic dip angle is less than 80°;† for stations with larger dip angles, extrapolation will provide additional tables if required.

The coefficients were computed by the method described by SCHMERLING (1957b, 1958), to which the present paper forms a sequel. The semi-graphical method originally described therein has been replaced by a machine programme which performs essentially the same steps. BUDDEN (1954) matrices (40×40) were used, covering the range 1·5 to 13·5 Mc/s with 0·3 Mc/s frequency increments.

2. REPRESENTATION OF THE EARTH'S MAGNETIC FIELD

The computation of the coefficients requires values of gyrofrequency and magnetic dip angle at ionospheric heights. Some method of specifying these must accordingly be found, both for the computations, and for determining the particular table to be used by any station.

The mean dipole approximation is rather crude, as widespread departures are found from a simple dipole field. The spherical harmonic expansion of VESTINE *et al.* (1947) is probably the best available, but a simpler method is suggested here which gives, very closely, the same results.

* The research reported in this paper has been sponsored by the USA National Committee for the IGY under Project No. 6·9 and by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF19(604)-1304.

† This is not an inherent limitation of the computations, but arises merely from a technicality in the computer programme. It is hoped to publish, at a later date, tables for larger values of dip angle.

VOL. 14 1959

If the ground value of the magnetic dip angle is used, together with an inverse-cube extrapolation from the ground field intensity, a separate dipole field is effectively taken for each point on the earth. That these values agree quite well with those given by VESTINE *et al.* may be seen from Table 1.

To compute the tables, the world was divided into eight equal zones of magnetic dip, at ten degree increments and ranging from 0° to 80°. Each dip-zone represents

Table 1. Comparison of inverse-cube and spherical harmonic fields

Long.	Colat.	Total field intensity (c.g.s. units) (from VESTINE)			Total field intensity (c.g.s. units) (from inverse-cube law)		Magnetic dip (from VESTINE)		
		h = 0 km;	h = 100 km;	h = 300 km	h ... 100 km;	h = 300 km	h = 0 km;	h = 100 km;	h = 300 km
300°	70°	0·440	0·420	0·381	0·420	0·383	53°41′	53°29′	53°16′
30°	30°	0·504	0·477	0·440	0·481	0·439	72°34′	72°24′	72°11′
90°	30°	599	0·567	0·514	0·571	0·522	77°17′	76°57′	76°29′
210°	50°	0·463	0·446	0·408	0·442	0·403	58°19′	58°57′	59°06′
270°	40°	0·0·609	0·576	0·522	0·581	0·531	79°02′	78°53′	78°33′

the dip for both positive and negative values of dip, e.g. the dip-zone for 10° to 20° dip represents the range in dip from +10° to +20° and from −10° to −20°. This is permissible in that the value of the coefficients depends only on the magnitude of the dip and not on the sign.

The variation of the total field in each zone was then examined, and two values chosen such that the extreme differences between the field at any place and the nearest chosen value of the field were minimized.* The tabular values of gyrofrequency were then obtained by an inverse-cube extrapolation to 200 km. By taking the value of the dip at the centre of each dip-zone and the two chosen values of gyrofrequency, two groups of coefficients were calculated to cover each dip-zone. Sixteen groups of coefficients were thus computed. The magnetic data were taken from the U.S. Navy Hydrographic Charts, published by the U.S. Navy Hydrographic Office.

3. ACCURACY OF THE COMPUTATIONS

The accuracy of the computations depends primarily on the frequency interval, Δf, used in the Budden matrix. For the special case of zero dip, the computed coefficients may be directly compared with the coefficients of KELSO (1952), which can be obtained from a table of sines to any desired accuracy. Table 2 gives the computed coefficients for $\Delta f = 0{\cdot}2$ Mc/s and 0·5 Mc/s together with the Kelso values. It may be seen that, even for $\Delta f = 0{\cdot}5$ Mc/s, agreement is to within a few parts per thousand for the largest coefficients. Considerable errors can be tolerated in the smaller coefficients, since the lowest portions of h'–f records usually have least curvature, resulting in only small changes of virtual height with frequency. It may thus be concluded that a choice of $\Delta f = 0{\cdot}3$ Mc/s leads to negligible errors from this cause.

* The values of gyrofrequency actually used in the computations are given in square brackets in the tables.

Table 2. Coefficients for zero dip

Kelso coefficients	Schmerling coefficients	
	$\Delta f = 0{\cdot}2$ Mc/s	$\Delta f = 0{\cdot}5$ Mc/s
0·997	0·997	0·997
0·972	0·973	0·974
0·924	0·924	0·925
0·853	0·852	0·852
0·760	0·760	0·758
0·649	0·648	0·646
0·522	0·521	0·519
0·383	0·384	0·378
0·233	0·237	0·229
0·078	0·081	0·078

4. Sensitivity to Magnetic Dip Angle and Gyrofrequency

Owing to the method used to calculate the tables of coefficients in this paper, there will be variations in the coefficients, for particular locations, from the coefficients designated, in the tables, for that station. It therefore becomes necessary to determine the sensitivity of the coefficients to dip and gyrofrequency in order to determine the maximum possible variation of the coefficients, presented in the tables, from the corresponding coefficients for a particular location.

The sensitivity of the coefficients to the magnetic parameters may be seen directly from the tables, but it is instructive to examine the effects of the two parameters separately.

Table 3 shows the variation in the coefficients due to a variation in dip angle of 20° (holding the gyrofrequency, f_H, constant at 1·18 Mc/s). Table 4 shows the variation in the coefficients due to a variation in gyrofrequency of 0·846 Mc/s (holding the dip angle constant at 60°).

Table 3. Variation of coefficients with dip angle

$f_H = 1{\cdot}18$ Mc/s Dip = 55°	$f_H = 1{\cdot}18$ Mc/s Dip = 35°
0·991	0·995
0·951	0·963
0·889	0·905
0·812	0·829
0·721	0·736
0·616	0·628
0·497	0·506
0·365	0·371
0·225	0·228
0·077	0·078

VOL. 14 1959

Now the maximum possible variation in dip and gyrofrequency of any station between magnetic dip angles of $\pm 80°$ from the dip and gyrofrequency of at least one table of coefficients, presented in this work, cannot exceed 5° and 0·225 Mc/s respectively. It can be seen from Table 3 that, for a variation in dip of 20° (holding f_H constant), the maximum variation for any of the coefficients is not greater than 2 per cent. Since the variation in the dip cannot exceed 5° the

Table 4. Variation of coefficients with gyrofrequency

$f_H = 0{\cdot}784$ Mc/s Dip = 60°	$f_H = 1{\cdot}62$ Mc/s Dip = 60°
0·991	0·989
0·952	0·943
0·891	0·875
0·815	0·794
0·724	0·702
0·618	0·598
0·498	0·481
0·366	0·353
0·222	0·216
0·075	0·073

variation in the coefficients, due to a 5° variation in dip, should not exceed $\frac{1}{2}$ per cent. Referring to Table 4, it can be seen that for a variation in gyrofrequency of 0·846 Mc/s (holding the dip constant) the variation for any of the coefficients is not greater than 3·2 per cent. The maximum possible variation in gyrofrequency envisaged is 0·225 Mc/s. This means that the variation in the coefficients due to a variation in gyrofrequency should not exceed 1 per cent.

Consequently, from Tables 3 and 4, it can be seen that the maximum variation in the first three coefficients (due to a variation in dip and gyrofrequency) of any location does not exceed $\frac{1}{2}$ per cent and the variation in the remaining coefficients does not exceed $1\frac{1}{2}$ per cent from the corresponding coefficients, given in the tables, for that particular location.

5. Overall Accuracy

The overall accuracy of record reduction is very difficult to specify, since this clearly depends on the exact nature of the record to be reduced. This discussion will therefore be carried out in terms of specific examples.

The reduction accuracy of single-layer h'–f records has already been considered by SCHMERLING (1957b, 1958). It was determined that the 5-point method gives results to within a few kilometres, except close to the critical frequency, where errors up to 10 km are possible. The 10-point method gives results to better than 5 km everywhere.

This discussion assumed good h'–f records, with flat lower boundaries. Scaling errors were not considered. If the lower boundary is not flat, errors due to low-frequency cut-off may be considerably greater, as shown by SCHMERLING (1957a).

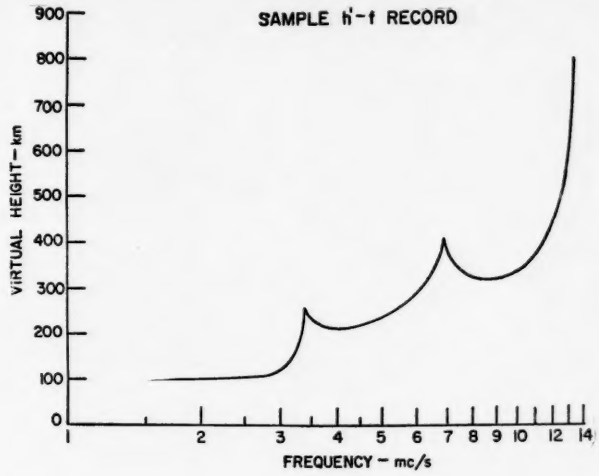


Fig. 1.

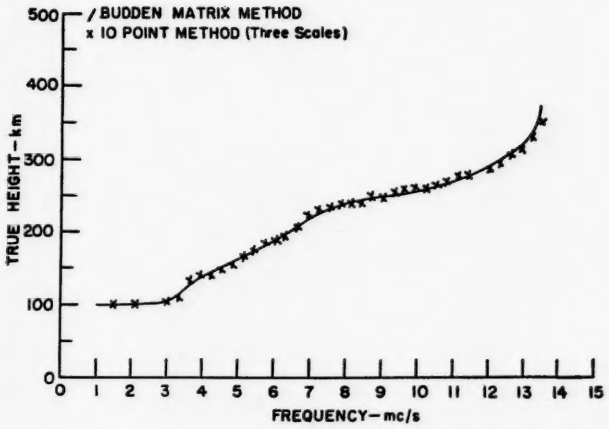


Fig. 2.

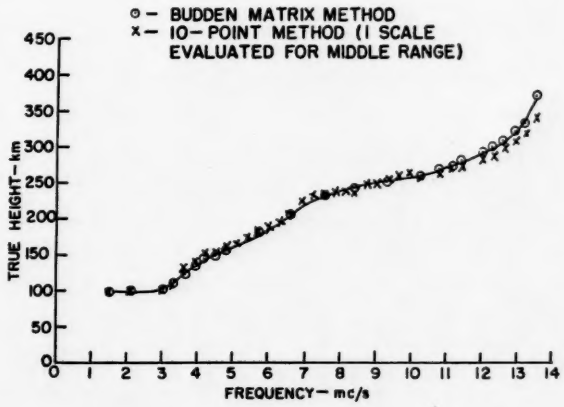


Fig. 3.

VOL.
14
1959

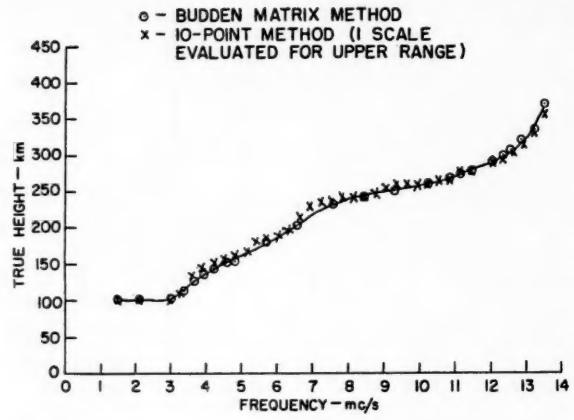


Fig. 4.

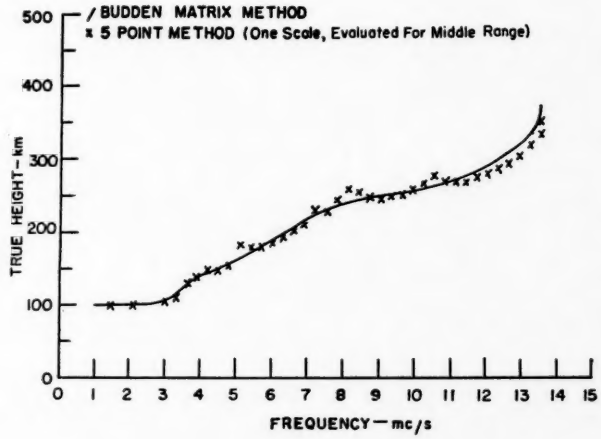


Fig. 5.

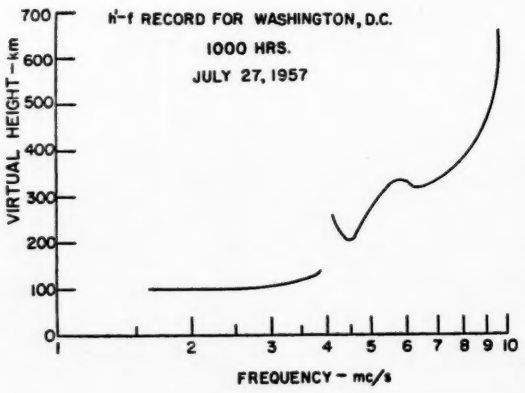


Fig. 6.

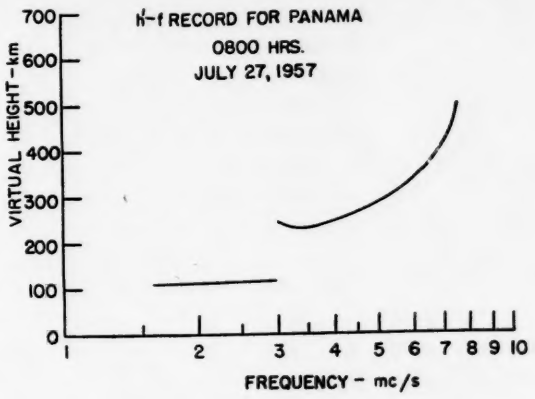


Fig. 7.

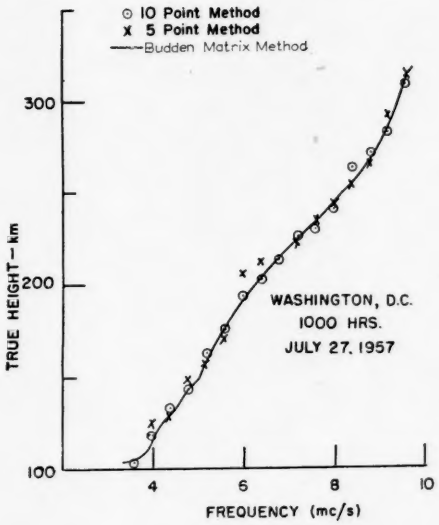


Fig. 8.

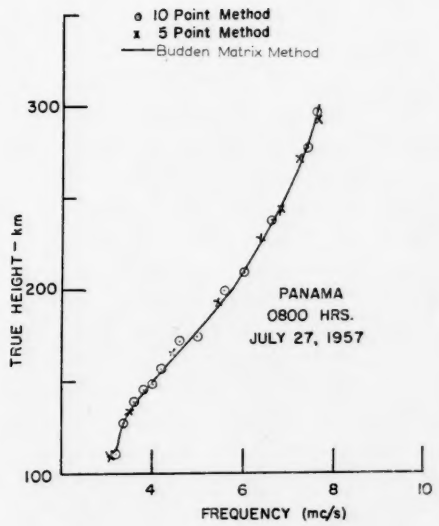


Fig. 9.

This effect is not peculiar to the present method of reduction, as it is inherent in the analysis, and will be found in any method. This is due to the fact that it arises from the necessity of inserting, by guesswork, a missing portion of the record.

To examine the errors which can arise in the reduction of more complicated h'–f records, an extreme example was taken. The h'–f record is shown in Fig. 1.

Fig. 2 shows the results of reduction by three 10-point scales (evaluated at the frequencies 3 Mc/s, 7·5 Mc/s and 13·5 Mc/s respectively), compared with the results of the Budden matrix method, which has been shown independently to be highly accurate. It can be seen that the agreement is to better than 5 km except for the last point, at the critical frequency, where scaling is very difficult. Fig. 3 shows the results of reduction using one 10-point scale, evaluated at the frequency 7·5 Mc/s. The results compare fairly well with those of the Budden method for frequencies up to 12 Mc/s. Above 12 Mc/s, use of this simple scale causes a large deviation from the Budden method. Fig. 4 shows the results of reduction using one 10-point scale, evaluated at the frequency 13·5 Mc/s. It can be seen that these results compare well with those of the Budden method except for the last point, at the critical frequency.

Fig. 5 shows the results of reduction using one 5-point scale, evaluated at the frequency 7·5 Mc/s. It can be seen that although extreme deviations of up to 20 km do occur the trend in the N–h profile is given quite well. The oscillations in both this and the previous curves are due to sampling errors, as discussed by Piggott (1954). They arise from intermediate cusps, and are clearly less severe for the 10-point reduction.

Two examples, taken from actual h'–f records, are shown in Figs. 6–9. Figs. 6 and 7 show the h'–f records and Figs. 8 and 9 the N–h profiles deduced. The solid line gives the profiles obtained from the Budden method, the crosses the results of 5-point scale analysis and the circles the results of the 10-point scale analysis. Agreement is seen to be quite close.

6. Tables of 10-Point and 5-Point Coefficients

The use of the following tables is best described by an example. Let us consider a station whose dip angle is −48° and whose gyrofrequency is 1·0 Mc/s. The coefficients for this station are found in the table whose range in dip is from 40° to 50°. There are two groups of coefficients corresponding to this dip-range: one group for the range in gyrofrequency from 0·580 Mc/s to 0·965 Mc/s, the other group for the range in gyrofrequency from 0·965 Mc/s to 1·35 Mc/s. The coefficients for the range in gyrofrequency from 0·965 Mc/s to 1·35 Mc/s are then used for this station. The upper columns are used in a 10-point analysis; the lower in a 5-point analysis.

Now, for each station there are eights sets of coefficients, each set corresponding to a particular frequency, f_n (f_n is the frequency at which the true height is desired). It might appear that one would need a set of coefficients for each frequency, f_n, in order to properly reduce an h'–f record. On the contrary, it has been shown in Section 5 that one set of coefficients used over a considerable frequency range

Dip (0°–10°)

	f_H (0·00–0·860) Mc/s [0·728 Mc/s]								f_H (0·860–1·12) Mc/s [1·00 Mc/s]							
	f_n (Mc/s)								f_n (Mc/s)							
	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997
	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973	0·973
	0·923	0·923	0·923	0·923	0·923	0·924	0·924	0·925	0·923	0·923	0·923	0·923	0·923	0·924	0·924	0·925
	0·852	0·852	0·851	0·851	0·851	0·851	0·851	0·851	0·852	0·852	0·851	0·851	0·851	0·851	0·851	0·851
	0·759	0·759	0·759	0·759	0·759	0·758	0·757	0·756	0·759	0·759	0·759	0·759	0·758	0·758	0·757	0·756
	0·648	0·648	0·648	0·648	0·647	0·647	0·646	0·643	0·648	0·648	0·648	0·648	0·647	0·647	0·646	0·643
	0·521	0·521	0·521	0·521	0·520	0·520	0·518	0·518	0·521	0·521	0·521	0·521	0·520	0·520	0·518	0·518
	0·382	0·381	0·381	0·381	0·380	0·380	0·378	0·383	0·381	0·381	0·381	0·381	0·380	0·380	0·378	0·383
	0·232	0·232	0·232	0·232	0·231	0·231	0·232	0·237	0·232	0·232	0·232	0·231	0·231	0·231	0·232	0·237
	0·078	0·077	0·077	0·077	0·078	0·078	0·079	0·082	0·077	0·077	0·077	0·077	0·078	0·078	0·079	0·082
$\frac{f}{f_n}$																
	0·988	0·988	0·988	0·988	0·989	0·989	0·989	0·988	0·988	0·988	0·988	0·988	0·989	0·989	0·989	0·988
	0·890	0·890	0·890	0·890	0·890	0·890	0·890	0·891	0·890	0·890	0·890	0·890	0·890	0·890	0·890	0·891
	0·706	0·706	0·706	0·705	0·705	0·705	0·704	0·702	0·706	0·706	0·706	0·705	0·705	0·705	0·704	0·702
	0·453	0·453	0·452	0·452	0·452	0·451	0·450	0·452	0·453	0·453	0·452	0·452	0·452	0·451	0·450	0·452
	0·155	0·155	0·155	0·155	0·155	0·155	0·156	0·161	0·155	0·155	0·155	0·155	0·155	0·155	0·156	0·161

Dip (10°–20°)

	f_H (0·580–0·855) Mc/s [0·728 Mc/s]								f_H (0·855–1·12) Mc/s [1·00 Mc/s]							
	f_n (Mc/s)								f_n (Mc/s)							
	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997	0·997
	0·971	0·971	0·971	0·972	0·972	0·972	0·972	0·971	0·971	0·971	0·971	0·972	0·972	0·972	0·972	0·971
	0·921	0·921	0·920	0·920	0·920	0·920	0·921	0·921	0·920	0·920	0·920	0·920	0·920	0·920	0·920	0·921
	0·849	0·848	0·848	0·848	0·847	0·847	0·846	0·845	0·848	0·848	0·847	0·847	0·846	0·846	0·845	0·845
	0·757	0·756	0·756	0·755	0·754	0·753	0·752	0·750	0·756	0·755	0·755	0·754	0·753	0·752	0·751	0·749
	0·646	0·646	0·645	0·645	0·644	0·643	0·641	0·637	0·645	0·645	0·644	0·644	0·643	0·642	0·640	0·637
	0·520	0·519	0·519	0·518	0·518	0·516	0·514	0·513	0·519	0·519	0·518	0·518	0·517	0·515	0·513	0·512
	0·381	0·380	0·380	0·379	0·379	0·377	0·376	0·379	0·380	0·380	0·379	0·379	0·378	0·377	0·375	0·378
	0·232	0·232	0·231	0·231	0·230	0·230	0·230	0·235	0·232	0·231	0·231	0·230	0·230	0·229	0·230	0·234
	0·077	0·077	0·077	0·077	0·077	0·078	0·078	0·081	0·077	0·077	0·077	0·077	0·077	0·077	0·078	0·080
$\frac{f}{f_n}$																
	0·988	0·988	0·988	0·988	0·988	0·988	0·987	0·987	0·987	0·987	0·988	0·988	0·988	0·988	0·988	0·987
	0·887	0·887	0·887	0·886	0·886	0·886	0·886	0·886	0·887	0·886	0·886	0·886	0·886	0·885	0·885	0·886
	0·704	0·703	0·703	0·702	0·701	0·700	0·698	0·696	0·703	0·702	0·702	0·701	0·700	0·699	0·698	0·695
	0·452	0·451	0·451	0·450	0·450	0·448	0·446	0·447	0·451	0·451	0·450	0·450	0·449	0·447	0·445	0·446
	0·155	0·155	0·154	0·154	0·154	0·154	0·155	0·159	0·155	0·155	0·154	0·154	0·154	0·154	0·155	0·159

Dip (20°–30°)

f_H (0·580–0·855) Mc/s [0·728 Mc/s]								f_H (0·855–1·15) Mc/s [1·00 Mc/s]							
f_n (Mc/s)								f_n (Mc/s)							
13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
0·996	0·996	0·996	0·996	0·996	0·996	0·996	0·995	0·996	0·996	0·996	0·996	0·996	0·996	0·996	0·995
0·969	0·969	0·969	0·969	0·969	0·969	0·968	0·967	0·968	0·968	0·968	0·969	0·969	0·969	0·968	0·967
0·917	0·917	0·916	0·916	0·916	0·915	0·915	0·915	0·916	0·916	0·915	0·915	0·914	0·914	0·914	0·914
0·846	0·845	0·844	0·843	0·842	0·841	0·839	0·837	0·844	0·843	0·842	0·841	0·840	0·839	0·837	0·836
0·754	0·753	0·752	0·751	0·750	0·748	0·745	0·741	0·752	0·751	0·750	0·749	0·747	0·745	0·743	0·739
0·644	0·643	0·642	0·641	0·640	0·638	0·635	0·630	0·642	0·642	0·641	0·639	0·638	0·635	0·632	0·627
0·518	0·518	0·517	0·516	0·515	0·513	0·510	0·507	0·517	0·516	0·516	0·515	0·513	0·511	0·508	0·504
0·380	0·379	0·379	0·378	0·377	0·375	0·373	0·374	0·379	0·378	0·378	0·377	0·376	0·374	0·371	0·372
0·231	0·231	0·231	0·230	0·229	0·229	0·229	0·232	0·231	0·231	0·230	0·229	0·229	0·228	0·228	0·230
0·077	0·077	0·077	0·077	0·077	0·077	0·078	0·080	0·077	0·077	0·077	0·077	0·077	0·077	0·077	0·079
$\frac{f}{f_n}$															
0·986	0·986	0·986	0·986	0·986	0·986	0·985	0·984	0·986	0·986	0·986	0·986	0·986	0·986	0·985	0·984
0·884	0·883	0·883	0·882	0·881	0·880	0·879	0·879	0·882	0·882	0·881	0·880	0·879	0·879	0·878	0·878
0·701	0·700	0·699	0·698	0·697	0·695	0·692	0·687	0·699	0·698	0·697	0·696	0·695	0·692	0·689	0·685
0·450	0·450	0·449	0·449	0·447	0·446	0·443	0·442	0·449	0·449	0·448	0·447	0·446	0·444	0·441	0·440
0·155	0·154	0·154	0·154	0·154	0·153	0·154	0·157	0·154	0·154	0·154	0·153	0·153	0·153	0·153	0·156

Dip (30°–40°)

f_H (0·580–0·900) Mc/s [0·756 Mc/s]								f_H (0·900–1·25) Mc/s [1·06 Mc/s]							
f_n (Mc/s)								f_n (Mc/s)							
13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
0·996	0·995	0·995	0·995	0·995	0·995	0·994	0·994	0·995	0·995	0·995	0·995	0·995	0·995	0·994	0·993
0·966	0·966	0·966	0·966	0·965	0·965	0·964	0·963	0·965	0·965	0·965	0·965	0·964	0·964	0·963	0·962
0·914	0·913	0·912	0·911	0·910	0·909	0·908	0·908	0·912	0·911	0·910	0·909	0·908	0·906	0·905	0·906
0·842	0·841	0·840	0·838	0·837	0·834	0·831	0·828	0·839	0·838	0·836	0·835	0·833	0·830	0·827	0·824
0·751	0·750	0·749	0·747	0·745	0·742	0·738	0·732	0·748	0·747	0·745	0·743	0·741	0·737	0·733	0·727
0·642	0·641	0·640	0·638	0·636	0·633	0·629	0·621	0·639	0·638	0·637	0·635	0·632	0·629	0·624	0·616
0·517	0·516	0·515	0·514	0·512	0·509	0·505	0·500	0·515	0·514	0·513	0·511	0·509	0·506	0·501	0·496
0·379	0·378	0·378	0·377	0·375	0·373	0·370	0·369	0·378	0·377	0·376	0·375	0·373	0·370	0·367	0·366
0·231	0·231	0·230	0·229	0·228	0·227	0·227	0·229	0·230	0·230	0·229	0·228	0·227	0·226	0·225	0·226
0·077	0·077	0·077	0·077	0·077	0·077	0·077	0·079	0·077	0·077	0·077	0·077	0·076	0·076	0·077	0·780
$\frac{f}{f_n}$															
0·984	0·984	0·984	0·984	0·984	0·983	0·982	0·981	0·984	0·984	0·984	0·983	0·983	0·983	0·982	0·981
0·880	0·880	0·879	0·877	0·876	0·874	0·872	0·870	0·878	0·877	0·875	0·874	0·872	0·870	0·868	0·867
0·698	0·698	0·696	0·695	0·692	0·689	0·685	0·678	0·696	0·694	0·693	0·691	0·688	0·685	0·680	0·673
0·449	0·449	0·448	0·447	0·445	0·443	0·439	0·436	0·448	0·447	0·446	0·444	0·442	0·440	0·435	0·432
0·154	0·154	0·154	0·153	0·153	0·153	0·153	0·155	0·154	0·154	0·153	0·153	0·152	0·152	0·152	0·153

Dip (40°–50°)

	f_H (0·580–0·965) Mc/s [0·784 Mc/s]								f_H (0·965–1·35) Mc/s [1·148 Mc/s]							
	f_n (Mc/s)								f_n (Mc/s)							
	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
	0·995	0·994	0·994	0·994	0·994	0·993	0·993	0·992	0·994	0·994	0·994	0·994	0·993	0·993	0·992	0·991
	0·963	0·963	0·962	0·962	0·962	0·961	0·960	0·957	0·961	0·960	0·960	0·960	0·959	0·959	0·958	0·955
	0·911	0·909	0·908	0·907	0·905	0·903	0·901	0·900	0·906	0·905	0·904	0·902	0·900	0·898	0·896	0·895
	0·839	0·838	0·836	0·834	0·831	0·828	0·823	0·818	0·834	0·832	0·830	0·828	0·825	0·821	0·816	0·811
	0·748	0·747	0·745	0·743	0·740	0·736	0·730	0·722	0·744	0·742	0·740	0·737	0·733	0·729	0·722	0·714
	0·640	0·639	0·637	0·635	0·632	0·628	0·622	0·613	0·636	0·634	0·632	0·630	0·626	0·621	0·615	0·604
	0·516	0·515	0·513	0·512	0·509	0·506	0·500	0·493	0·512	0·511	0·510	0·507	0·504	0·500	0·494	0·486
	0·378	0·377	0·376	0·375	0·373	0·371	0·366	0·364	0·376	0·375	0·374	0·372	0·370	0·367	0·362	0·358
	0·231	0·230	0·229	0·228	0·227	0·226	0·225	0·226	0·230	0·229	0·228	0·227	0·225	0·224	0·222	0·222
	0·077	0·077	0·077	0·077	0·076	0·076	0·077	0·078	0·077	0·076	0·076	0·076	0·076	0·076	0·076	0·076
$\frac{f}{f_n}$																
	0·982	0·982	0·982	0·981	0·981	0·980	0·979	0·977	0·981	0·981	0·981	0·980	0·980	0·979	0·978	0·976
	0·877	0·876	0·874	0·873	0·870	0·867	0·864	0·861	0·872	0·871	0·869	0·867	0·864	0·861	0·858	0·854
	0·696	0·695	0·693	0·691	0·688	0·684	0·678	0·669	0·692	0·690	0·688	0·685	0·682	0·677	0·670	0·660
	0·448	0·447	0·446	0·445	0·443	0·440	0·434	0·430	0·446	0·445	0·443	0·441	0·439	0·435	0·429	0·423
	0·154	0·154	0·153	0·153	0·152	0·152	0·152	0·153	0·153	0·153	0·152	0·152	0·151	0·150	0·150	0·150

VOL. 14 1959

Dip (50°–60°)

	f_H (0·70–1·10) Mc/s [0·896 Mc/s]								f_H (1·10–1·50) Mc/s [1·316 Mc/s]							
	f_n (Mc/s)								f_n (Mc/s)							
	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
	0·993	0·993	0·993	0·993	0·992	0·992	0·991	0·990	0·993	0·992	0·992	0·992	0·991	0·991	0·990	0·989
	0·960	0·959	0·958	0·958	0·957	0·956	0·954	0·951	0·956	0·955	0·955	0·954	0·953	0·952	0·951	0·948
	0·906	0·904	0·903	0·901	0·898	0·895	0·892	0·889	0·900	0·898	0·897	0·894	0·892	0·888	0·885	0·882
	0·834	0·832	0·830	0·827	0·824	0·819	0·813	0·806	0·828	0·825	0·823	0·819	0·815	0·810	0·804	0·796
	0·744	0·742	0·740	0·737	0·733	0·728	0·720	0·709	0·738	0·736	0·733	0·729	0·724	0·718	0·710	0·699
	0·636	0·635	0·633	0·630	0·626	0·621	0·613	0·601	0·631	0·629	0·626	0·623	0·618	0·612	0·604	0·590
	0·513	0·512	0·510	0·508	0·505	0·500	0·493	0·483	0·509	0·508	0·505	0·502	0·499	0·493	0·485	0·473
	0·376	0·375	0·374	0·372	0·370	0·366	0·361	0·356	0·374	0·373	0·371	0·369	0·366	0·362	0·355	0·348
	0·230	0·229	0·228	0·227	0·225	0·223	0·221	0·220	0·228	0·228	0·227	0·225	0·223	0·221	0·218	0·215
	0·076	0·076	0·076	0·076	0·076	0·076	0·076	0·076	0·076	0·076	0·076	0·075	0·075	0·075	0·074	0·074
$\frac{f}{f_n}$																
	0·980	0·980	0·979	0·979	0·978	0·977	0·976	0·973	0·978	0·978	0·977	0·977	0·976	0·975	0·974	0·971
	0·872	0·871	0·869	0·866	0·863	0·859	0·855	0·849	0·866	0·864	0·862	0·859	0·855	0·851	0·846	0·841
	0·692	0·691	0·688	0·685	0·682	0·676	0·669	0·656	0·687	0·684	0·681	0·678	0·673	0·667	0·658	0·645
	0·446	0·445	0·444	0·442	0·439	0·435	0·428	0·421	0·443	0·442	0·440	0·437	0·434	0·429	0·421	0·412
	0·153	0·153	0·152	0·151	0·151	0·150	0·149	0·149	0·153	0·152	0·151	0·150	0·149	0·148	0·147	0·146

Dip (60°–70°)

f_H (0·780–1·20) Mc/s [0·980 Mc/s]								f_H (1·20–1·65) Mc/s [1·428 Mc/s]							
f_n (Mc/s)								f_n (Mc/s)							
13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
0·992	0·992	0·992	0·991	0·991	0·990	0·989	0·988	0·991	0·991	0·990	0·990	0·990	0·989	0·988	0·986
0·956	0·955	0·954	0·953	0·952	0·951	0·949	0·946	0·951	0·950	0·949	0·948	0·947	0·946	0·944	0·941
0·901	0·899	0·897	0·894	0·891	0·888	0·883	0·879	0·894	0·892	0·890	0·887	0·883	0·879	0·875	0·870
0·829	0·826	0·824	0·820	0·816	0·810	0·803	0·793	0·822	0·819	0·815	0·811	0·806	0·800	0·792	0·782
0·739	0·736	0·734	0·730	0·725	0·719	0·710	0·696	0·732	0·729	0·726	0·721	0·715	0·708	0·698	0·683
0·632	0·629	0·627	0·623	0·619	0·612	0·603	0·588	0·626	0·624	0·620	0·616	0·610	0·603	0·592	0·576
0·509	0·507	0·505	0·502	0·498	0·492	0·484	0·472	0·505	0·503	0·500	0·497	0·492	0·485	0·474	0·461
0·373	0·372	0·370	0·368	0·365	0·360	0·353	0·347	0·371	0·369	0·367	0·365	0·361	0·355	0·347	0·338
0·227	0·226	0·225	0·223	0·221	0·219	0·217	0·214	0·226	0·225	0·224	0·222	0·219	0·216	0·213	0·208
0·075	0·075	0·075	0·074	0·074	0·074	0·074	0·073	0·075	0·075	0·074	0·074	0·074	0·073	0·072	0·071
$\frac{f}{f_n}$															
0·977	0·977	0·976	0·976	0·975	0·974	0·972	0·969	0·974	0·974	0·974	0·973	0·972	0·971	0·969	0·966
0·867	0·865	0·862	0·859	0·855	0·851	0·845	0·838	0·860	0·857	0·854	0·851	0·846	0·841	0·835	0·827
0·687	0·685	0·682	0·679	0·674	0·667	0·658	0·643	0·681	0·678	0·675	0·670	0·665	0·657	0·646	0·630
0·442	0·441	0·439	0·436	0·433	0·428	0·419	0·410	0·440	0·438	0·435	0·432	0·428	0·421	0·411	0·400
0·151	0·150	0·150	0·149	0·148	0·147	0·146	0·145	0·151	0·150	0·149	0·148	0·147	0·145	0·143	0·141

Dip (70°–80°)

f_H (1·25–1·55) Mc/s [1·40 Mc/s]								f_H (1·55–1·85) Mc/s [1·70 Mc/s]							
f_n (Mc/s)								f_n (Mc/s)							
13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0	13·5	12·0	10·5	9·0	7·5	6·0	4·5	3·0
0·990	0·990	0·989	0·989	0·988	0·988	0·987	0·985	0·988	0·988	0·988	0·987	0·987	0·986	0·985	0·983
0·948	0·947	0·945	0·944	0·943	0·941	0·940	0·937	0·944	0·943	0·941	0·940	0·938	0·937	0·935	0·932
0·890	0·888	0·885	0·882	0·878	0·874	0·869	0·864	0·884	0·881	0·878	0·874	0·870	0·866	0·860	0·856
0·817	0·814	0·810	0·806	0·801	0·794	0·786	0·776	0·810	0·806	0·802	0·797	0·792	0·784	0·775	0·765
0·728	0·725	0·721	0·716	0·710	0·702	0·692	0·698	0·720	0·716	0·712	0·707	0·700	0·692	0·680	0·666
0·622	0·619	0·616	0·611	0·605	0·598	0·587	0·573	0·615	0·612	0·607	0·602	0·596	0·587	0·575	0·559
0·501	0·499	0·496	0·492	0·488	0·481	0·471	0·459	0·495	0·492	0·489	0·484	0·479	0·471	0·459	0·445
0·368	0·367	0·364	0·361	0·358	0·352	0·345	0·336	0·364	0·361	0·359	0·355	0·352	0·344	0·335	0·326
0·224	0·223	0·222	0·220	0·218	0·216	0·211	0·207	0·222	0·220	0·219	0·216	0·213	0·210	0·206	0·201
0·075	0·075	0·074	0·074	0·074	0·073	0·073	0·071	0·074	0·073	0·073	0·073	0·072	0·071	0·070	0·069
$\frac{f}{f_n}$															
0·972	0·971	0·971	0·970	0·969	0·968	0·966	0·963	0·969	0·969	0·969	0·968	0·967	0·966	0·964	0·961
0·856	0·853	0·850	0·847	0·841	0·835	0·829	0·821	0·848	0·845	0·842	0·838	0·832	0·826	0·819	0·811
0·677	0·674	0·670	0·665	0·659	0·651	0·641	0·626	0·669	0·666	0·662	0·656	0·649	0·641	0·629	0·613
0·436	0·434	0·432	0·428	0·424	0·418	0·408	0·398	0·431	0·429	0·426	0·422	0·417	0·417	0·399	0·387
0·150	0·149	0·149	0·148	0·147	0·145	0·143	0·139	0·147	0·146	0·145	0·144	0·143	0·141	0·138	0·135

yields very good results. If it is desired to obtain the coefficients for a frequency not given in the tables one can interpolate to obtain the desired coefficients.

The method of use of the coefficients, for the actual reduction of h'–f records, is described fully by SCHMERLING (1957b, 1958).

7. CONCLUSIONS

Tables of coefficients for the rapid reduction, without computing aids, of h'–f records to N–h profiles have been presented for any station whose magnetic dip angle does not exceed 80°.

It has been shown that single-layer h'–f records are adequately reduced by the 5-point method, except very close to the critical frequency. The 10-point method should be used in all other cases. If one 10-point scale is to be used, it should be evaluated for the high frequency end of the record. Using three 10-point scales, results to better than 5 km in true height are easily attainable, even for the most complex h'–f records found in practice.

REFERENCES

BUDDEN K. G.	1954	*Cambridge Conference on Ionospheric Physics* p. 332. Physical Society, London.
KELSO J. M.	1952	*J. Geophys. Res.* **57,** 357.
PIGGOTT W. R.	1954	*J. Atmosph. Terr. Phys.* **5,** 201.
SCHMERLING E. R.	1957a	*Science Report No.* 94. Ionosphere Research Laboratory, Pennsylvania State University.
SCHMERLING E. R.	1957b	*Science Report No.* 96. Ionosphere Research Laboratory, Pennsylvania State University.
SCHMERLING E. R.	1958	*J. Amosph. Terr. Phys.* **12,** 8.
VESTINE *et al.*	1948	*Publication* 578, Carnegie Institute.

VOL. 14 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 262 to 272. Pergamon Press Ltd. Printed in Northern Ireland

Interpretation of smooth type atmospheric waveforms

F. HEPBURN
Department of Physics, University of Nottingham

(*Received* 25 *July* 1958)

Abstract—Following a brief survey of the conflicting opinions in this field and general remarks on pulse reflection and waveguide interpretations of atmospheric waveforms, the analyses of smooth type waveforms of varying degrees of irregularity are considered. Since observation suggests such irregularities are a source phenomenon, the implications of various discharge currents in the lightning channel are discussed in terms of waveguide propagation in the first-order mode and are seen to explain regular and pseudo regular waveforms. The simple reflection theory is also extended to more complicated source conditions to give analogous conclusions. Obviously irregular waveforms are tentatively ascribed to discharges of pulsating character.

Amplitude analysis of the waveform oscillation is shown to give the waveform spectrum very simply so that observations of a statistical nature can yield both the average spectrum at the source and the relative attenuation coefficients for the respective frequencies which are propagated.

A review of previous work in the light of the present paper is thought to resolve the apparent discrepancies.

1. INTRODUCTION

THE early observation of the damped-sinusoidal nature of the day-time electric field change waveforms due to atmospheric disturbances (APPLETON and CHAPMAN, 1937) was initially considered as evidence in favour of an oscillatory discharge in the lightning channel (LUTKIN, 1939). Despite the original suggestion by LABY *et al.* (1938) that the oscillatory structure was due to propagation, they were later (1940) of the opinion that the propagation mechanism alone was inadequate to explain all of the waveform and concluded that some oscillation did occur in the lightning channel. However, SCHONLAND *et al.* (1940) were satisfied by a direct application of reflection theory to day-time waveforms.

This and later discussion of atmospheric waveforms has been complicated by the absence of an acceptable comprehensive classification, such as that indicated recently (HEPBURN, 1957) and later described more fully (1958), so that it is uncertain when various workers have been considering the same phenomena. HALES (1948) has suggested that the smooth waveforms observed by day in South Africa are adequately explained in terms of first-order mode propagation between perfectly conducting earth and ionosphere boundaries. The success of the neglect of other order propagation is due to the imperfectly conducting nature of the boundaries, which does not greatly change first-order propagation over the relevant frequency band but modifies the other order responses to a large extent (BUDDEN 1951; LIEBERMANN 1956).

CATON and PIERCE (1952) state that day-time waveforms and the quasi-sinusoidal night-time variety of their Regular Smooth Type are generally inexplicable in terms of the reflection theory with waveform crests corresponding to ionosphere reflected pulses and suggest they may be explicable in terms of pulse propagation between imperfectly conducting boundaries. However, HORNER

and CLARKE (1955) have successfully analysed distant waveforms from the south-west in terms of the simple reflection theory.

Subsequent investigation suggests that HALES' criterion of success was not stringent and his conclusions somewhat optimistic, although in broad outline the concept appears acceptable. The investigations of HORNER and CLARKE have shown that practical analysis of waveforms showing only a few oscillations is incapable of giving reflection height or pulse order unambiguously so it is surprising that CATON and PIERCE (1952) were unable to obtain acceptable analyses from at

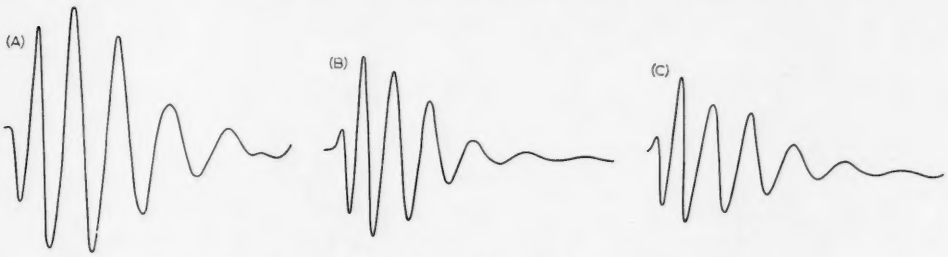


Fig. 1. Regular, pseudo regular and irregular smooth type waveforms.

least some of the waveforms of this type. Their implication of different conclusions from pulse reflection and waveguide approaches requires further examination to discover why, if at all, two differing conclusions could arise.

The present paper is concerned with the interpretation of smooth type return stroke waveforms (HEPBURN, 1958) which are recorded by day and night from storm centres to the west and south-west of the British Isles at distances greater than 1700 km using a conventional wide frequency band recording system. These waveforms consist of about half-a-dozen smooth oscillations, whose periods increase gradually from less than 80 to 250 μsec throughout the record and whose amplitudes increase rapidly to a maximum during the first third of the trace, thereafter decreasing steadily to zero (Fig. 1). Many of the waveforms show an obvious irregularity in the simple period and amplitude variations along the trace, which has formed the basis of subdivision of this type of waveform.

2. Pulse Reflection and Wave-guide Interpretation

In a previous paper (HEPBURN, 1957) it was shown that, for simple pulses and perfectly conducting boundaries, the result of combining the various waveguide mode responses was identical in form and time relations with the response considered as specular reflection. For the case of imperfect boundaries in which the first-order mode response predominates, complications were shown to arise when the initial pulse had appreciable duration resulting in distortion of the reflected pulses and invalidity of the assumption that the crests of the waveform were appropriate points for comparison with pulse reflection analysis. In the case of waveforms of almost pure first-order response (i.e. very distant long train or smooth types) the reflected pulses appear as sinusoidal oscillations and any distortion occurs principally as displacement of the crests along the waveform. For simple single or double primary pulses there is no relative displacement, but

the reflection theory formula is applicable to crests of the oscillation in the former case and zeros in the latter. However, it is extremely doubtful if any observer would not obtain a good analysis because of measuring crests instead of zeros on a simple smooth waveform, equivalent to an error of only one-quarter in the assumed pulse order. It would thus appear that any possible different conclusion derived from waveguide considerations is not essentially due to the difference of approach but rather to over-simplification of the complicated source conditions in one case. It will be shown later that a considerable proportion of smooth type waveforms may be interpreted in terms of waveguide nomenclature, because this approach is more amenable to extension to complicated source phenomena than is convenient with the pulse reflection method. This is especially so in these cases of imperfectly conducting boundaries when the propagation is almost entirely in the first-order mode.

3. General Analysis of Smooth Type Waveforms

Waveguide analysis has certain advantages in the interpretation of smooth type atmospherics since it avoids the assumption that one particular feature of the waveform, usually the positive or negative peak, should conform to some derived relation as is the case with simple reflection theory analysis. Following the method of HALES, if the quasi-period τ at a given point along the waveform is plotted as a function of the time of occurrence T of the point reckoned from the start of the trace, the resulting plot is intermediate between two limits corresponding to regular and extremely irregular varieties of waveform. Frequently the analysis of pseudo-regular waveforms strongly suggests the existence of a common minor irregularity discernable above the random errors of measurement. Examples of waveforms and their analyses are given in Figs. 1 and 2. The estimates of period were obtained from the durations of half-cycles of the oscillation measured between alternate positive and negative crests (○) or between successive intersections of the axis and the waveform trace (●). The smooth theoretical curves are for the specified values of ionosphere height h and storm distance d.

Analyses from extremely regular waveforms (Fig. 2A) are consistent with first-order mode propagation of sharp initial pulses but the short duration of the traces allows a considerable range of solutions for the propagation distance and reflection height (cf. HALES, 1948; CATON and PIERCE, 1952; HORNER and CLARKE, 1955), so that neither can be determined significantly in the absence of further information. Less regular waveforms cannot be interpreted so simply.

As regular waveforms and traces showing various degrees of irregularity are recorded from the same source region during a single recording period, the irregularities are more reasonably attributed to source phenomena than variations in propagation. An approach to this problem has been made (HEPBURN, 1957) when it was shown that plausible variations of current in the lightning channel can radiate complex signals, of varying amplitude spectra, the phases of whose components are related in various simple ways to their respective frequencies. These were expressed in terms of the Fourier transform of the source signal $A(f) \cdot \exp 2\pi\theta(f)i$, where $A(f)$ gives the amplitude of a wave of given frequency which is radiated and $\theta(f)$ its initial phase. Under certain conditions $\theta(f)$ is a linear

function of frequency over the appropriate range. In another case $\theta(f)$ is approximately represented by two different linear functions applicable to ranges above and below a transition frequency. It is the aim of the present paper to show that the apparently regular waveforms may be interpreted in terms of such source conditions.

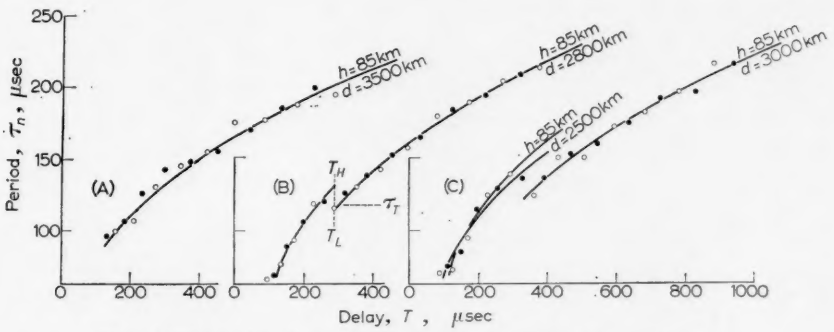


Fig. 2. Waveguide theory analyses of the respective waveforms.
Taking $h = 85$ km, waveform A gives $d = 3500 \pm 200$ km; B gives $d = 2800 \pm 200$ km, $\dot{\theta}_H - \dot{\theta}_L = 60$ μsec, $\tau_T \simeq 117$ μsec; C gives $d = 3000 \pm 500$ km, $\dot{\theta}_H - \dot{\theta}_L \simeq 90$ μsec(?), $\tau_T \simeq 135$ μsec.

4. Waveguide Interpretation of Apparently Regular Waveforms

Using the nomenclature of the previous paper (Hepburn, 1957), for a general source and first-order mode propagation, a feature of the form

$$\cos 2\pi\{f_n(t - t_n) + \theta_n + \phi_n\}$$

where $t \simeq t_n$ occurs at a time

$$t_n = d/c - \dot{\theta}(f_n) - \dot{\phi}(f_n) \tag{1}$$

This may be written as

$$t_n = d/u = d(1 - \tau^2c^2/4h^2)^{-\frac{1}{2}}/c - \dot{\theta}(f_n) \tag{2}$$

showing that the delay $T = t_n - t_0$ of a given signal of period τ with respect to the start of the waveform at t_0 conforms to simple wave guide dispersion so long as $\dot{\theta}(f_n)$ is constant. Such a truly regular waveform and its analysis are shown in Figs. 1A and 2A. It is of interest to investigate the predictions of the source analysis in the more complicated case. When $\dot{\theta}(f_L)$, $\dot{\phi}(f_L)$, $\dot{\theta}(f_H)$, $\dot{\phi}(f_H)$ are the slopes of the source and propagation characteristics in the low and high frequency regimes and θ_L, ϕ_L, θ_H, ϕ_H are the corresponding intercepts—only the θ terms being constants—the frequencies within each regime separately conform to the waveguide dispersion formula but the low frequency portion of the waveform is retarded by $(\dot{\theta}_H - \dot{\theta}_L)$ relative to the higher frequencies. In the analysis of such a pseudo regular waveform (Fig. 1B) the (T, τ) points would lie either on the initial portion of the curve given by:

$$T = d\{(1 - \tau^2c^2/4h^2)^{-\frac{1}{2}} - 1\}/c \tag{3}$$

or the later portion of the curve:

$$T = d\{(1 - \tau^2c^2/4h^2)^{-\frac{1}{2}} - 1\}/c + (\dot{\theta}_H - \dot{\theta}_L) \tag{4}$$

Fig. 2B indicates the extent to which this is shown in practice. By obtaining the period from only one half-cycle of oscillation and at intervals of only one quarter-cycle along the trace a reasonable estimate may be attempted of the transition frequency f_T and hence the initial phase difference at the source from the relation:

$$(\theta_L - \theta_H) = f_T(\dot{\theta}_H - \dot{\theta}_L) \tag{5}$$

As the low frequency portion of the waveform is retarded, it is at first surprising to notice that the analytical points in the transition region do not lie above the simple theoretical curves corresponding to abnormally extended periods. However in this transition region there is a phase change necessitating the inclusion of a considerable fraction of a cycle of oscillation, namely $(\theta_H - \theta_L)$, which is accommodated by the low frequency retardation. Considering the crests of the waveform, the condition for the nth crest is $\theta_n + \phi_n = n$. Suppose the transition between the high and low frequency regimes occurs in the vicinity of the nth peak. By extrapolation through the transition point from the earlier part of the waveform containing the higher frequencies, the nth crest occurs when $\theta_H + \phi_H = n$ at $t_H = d/c - \dot{\theta}(f_H) - \dot{\phi}(f_H)$ (see Fig. 2B). Extrapolation from the lower frequency later portions of the waveform would give the nth crest when $\theta_L + \phi_L = n$ at $t_L = d/c - \dot{\theta}(f_L) - \dot{\phi}(f_L)$. Since $t_L = t_H$, no discontinuity in the phase of oscillation in the waveform is expected. However the quasi-period of the crest by extrapolation from the higher frequencies is significantly greater than that corresponding to the lower frequencies (see Figs. 2B, 3B), so that in passing through the transition region the quasi-period increases less rapidly than in the simple case.

VOL
14
195

5. Reflection Theory Interpretation of Apparently Regular Waveforms

It is instructive to investigate the pulse reflection aspect to compare the results expected from such analysis of smooth waveforms with those from waveguide considerations. The occurrence of a feature in the first-order mode response of the form $\cos 2\pi\{f_n(t - t_n) + \theta_n + \phi_n\}$ when $t \simeq t_n$ is given by:

$$t_n = (4h^2\phi^2{}_n/c^2 + d^2/c^2)^{\frac{1}{2}} - \dot{\theta}(f_n) \tag{6}$$

showing that identity with reflection theory relative time intervals is obtained when $\dot{\theta}(f_n)$ is constant and ϕ_n integral. When a sharp single pulse is radiated θ_n is zero so that crests of the waveform satisfy the relation. A graph of crest separation (i.e. quasi period) $\tau'_n = t_n - t_{n-1}$ against crest reflection order n will be of the predicted form and have an intercept n_0 on the order axis of the form $(N+\frac{1}{2})$, which will be $\frac{1}{2}$ when no mistake has been made in ascribing the crest orders. This is so because of the approximate relation, valid for small n, that $\tau'_n \simeq 2h^2(2n - 1)/cd$. Slight confusion is introduced here by retaining conformity with established nomenclature, which defines quasi period τ_n as $t_{n+\frac{1}{2}} - t_{n-\frac{1}{2}}$ in waveguide treatment whereas the analogous quantity, crest separation, τ'_n for pulse reflection is taken as $t_n \sim t_{n\pm1}$. This difference accounts for the non-zero intercept on the order axis in the case of (n, τ') plots. If the negative crests of the waveform are measured or the discharge is of the opposite polarity the intercept will be integral valued. If the radiated signal is a double impulse, $\theta_n = \frac{1}{4}$ and the zeros following the crests

are the points whose time relations satisfy the formula for integral ϕ_n. The time separations of the crests lie on the same (n, τ') graph but for non-integral peak order values ϕ_n of the form $(n \pm \frac{1}{4})$ (cf. Fig. 3A).

Reflection theory analysis of a waveform possessing the more complicated source conditions considered above results in a (n, τ') graph having a discontinuity where the crest separation becomes equal to the transition period, Fig. 3B. In so far as $\theta_H = 0$, the earlier crests of the waveform satisfy reflection theory

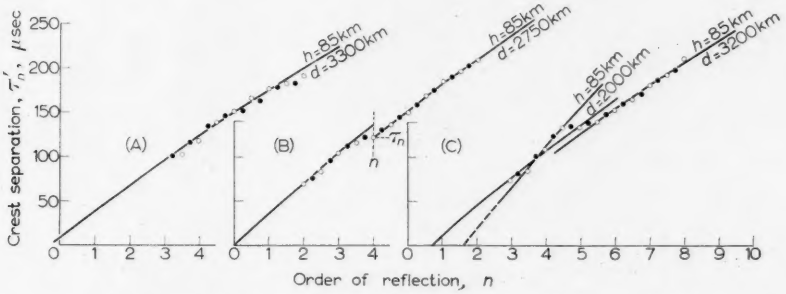


Fig. 3. Reflection theory analyses of the respective waveforms.
Taking $h = 85$ km, waveform A gives $d = 3300 \pm 200$ km, $n_0 = -\frac{1}{4} \pm \frac{1}{4}$; B gives $d = 2750 \pm 300$ km, $\dot{\theta}_L - \dot{\theta}_H \simeq \frac{1}{2}c$, $\tau_T = 122$ μsec, C gives $d = 3200 \pm 200$ km, $\dot{\theta}_L - \dot{\theta}_H \simeq \frac{1}{4}c(?)$, $\tau_T = 135$ μsec.

predictions. However for the later lower frequency oscillations $\theta_L \simeq \frac{1}{4}$ and alternate zeros of the trace satisfy the formula for successive integral values of ϕ_n. The crests preceding these alternate zero values will lie on the same graph for non-integral peak order values of the form $(n - \frac{1}{4})$. When measuring crests throughout the waveform, the (n, τ') points from the later portion of the trace lie on a curve of the same form as that through the earlier points but displaced by a quarter of an order, or, more generally, by $(\theta_L - \theta_H)$. Thus the discontinuity in the source characteristic results in a crest separation–crest order curve with a break when the crest separation equals the transition period τ_T and relative displacement of the two portions amounts to $(\theta_L - \theta_H)$ of an order. The break in the analogous quasi-period-delay curve occurs for the same condition and the relative displacement is:

$$(\dot{\theta}_H - \dot{\theta}_L) = \tau_T(\theta_L - \theta_H) \tag{7}$$

Unfortunately deviations from the simple case whereby this occurrence may be detected are comparable with the experimental errors in determining the time intervals from the waveforms so that it is difficult to be certain in ascribing the deviations to this cause. Peak positions may be known to an accuracy of ± 1 μsec at the beginning of the trace but this falls to ± 5 μsec at best as the oscillation amplitude decreases. A phase change of a quarter-cycle in a waveform from 3000 km, for which the peak separation increases with peak order at the rate of about 30 μsec/cycle, would give portions of the (n, τ') curve displaced vertically by only 7 μsec in time.

6. Pseudo Regular Waveforms

Analysis of pseudo regular waveforms on the basis of a waveguide or pulse reflection theory extended to include the more complicated source conditions discussed in Sections 4 and 5 would appear superficially to be quite acceptable. However in the particular case considered in detail, it is surprising to find the source phase discontinuity ($\theta_L - \theta_H$) considerably in excess of a quarter-cycle. It is difficult to understand how such phase discontinuities could arise unless the lightning discharge current is of a pulsating character. Because of limited accuracy

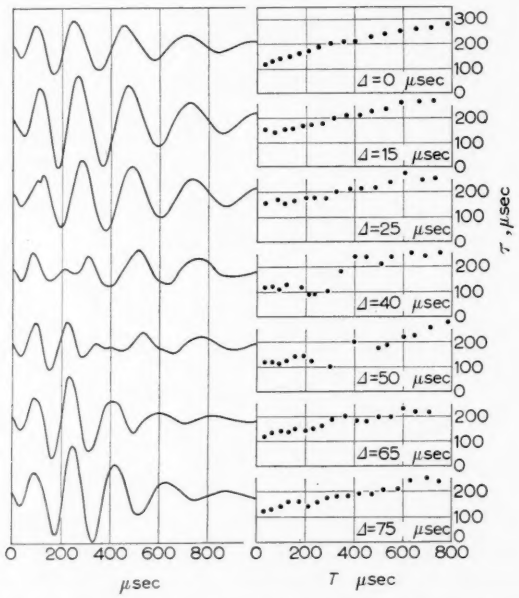


Fig. 4. Waveforms and waveguide analyses for a discharge current consisting of equal pulses separated by the specified intervals.

in the determination of points on the waveforms, too much significance must not be placed upon these estimates as the theoretical curves for slightly different storm distances would give values of only a quarter-cycle. The curves fitted to the points gave the best agreement without reference to this subsidiary requirement of an intelligible value for the phase discontinuity.

7. Irregular Waveforms

Analysis of irregular waveforms shows that they cannot be interpreted in terms of the modified source conditions already considered. Occasionally the analytical points from only slightly irregular waveforms do not lie on displaced portions of a single simple theoretical curve and the range estimates from the beginning and end portions of the waveform are incompatible (Figs. 2C and 3C). Such traces are of far too frequent occurrence to be explained in terms of overlapping waveforms from simultaneous discharges at different distances.

It is known from the work of Malan and Collens (1937) that return strokes

may be pulsating in character, due to the recrudescence of current when the return stroke reaches a major branching point of the leader discharge. As an exercise in the implications of this knowledge, computations have been made of the waveforms resulting from the occurrence of two simple identical discharges separated by various time delays. The resulting analyses emphasize the irregularities of these waveforms (Fig. 4) even when cursory inspection gives little indication of extreme irregularity. Although in practice the components of such a complex discharge are

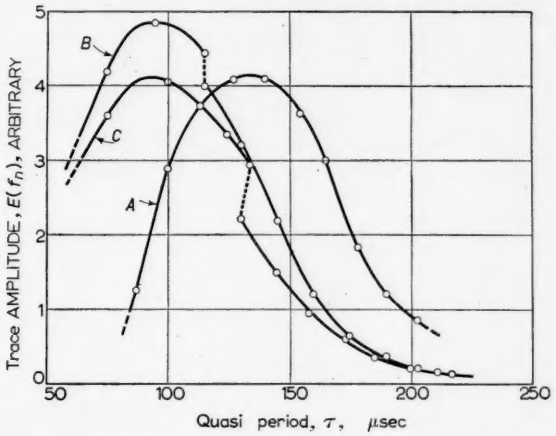


Fig. 5. Variations of trace height with quasi-period for the respective waveforms.

unlikely to be identical, the exercise clearly illustrates the potentialities of the mechanism for producing a wide range of waveform irregularities. In view of the known occurrence of such discharges, it appears to be reasonable and acceptable to attribute irregular waveforms to this process.

8. Amplitude Analysis

The Fourier component amplitudes in the received signal are easily derived from the variation of trace height with quasi-period (Fig. 5) for the regular waveforms which may be directly interpreted in terms of first-order mode propagation. In general allowing for two dimensional divergence of the radiation and attenuation in propagation, the mode response at a time t_n is of the form:

$$E(t, d, m) = A(f_n) \cdot d^{-\frac{1}{2}} \cdot \exp(-\alpha_f \cdot d) \{\ddot{\theta}(f_n) + \ddot{\phi}(f_n)\}^{-\frac{1}{2}} \cdot \cos 2\pi\{f_n(t - t_n) + \theta_n + \phi_n\} \quad (8)$$

$$= E(f_n) \cdot \cos 2\pi\{f_n(t - t_n) + \theta_n + \phi_n\} \quad (9)$$

where $t \simeq t_n$

The amplitude spectrum of the received signal $S(f_n)$ for a particular waveform is given in terms of the trace amplitude $E(f_n)$ by:

$$S(f_n) = E(f_n) \cdot \{\ddot{\theta}(f_n) + \ddot{\phi}(f_n)\}^{\frac{1}{2}} \quad (10)$$

The phase variation term $\ddot{\theta}(f_n)$ due to source conditions is zero, when the radiated pulse is very short, and very small at frequencies other than the transition

frequency f_T for the more complex source characteristic. The phase term $\ddot{\phi}(f_n)$ due to propagation (see Fig. 6) is given by:

$$\ddot{\phi}(f_n) = d \,.\, f_c{}^2/c \,.\, (f^2 - f_c{}^2)^{3/2} \tag{11}$$

so that the received signal has the spectrum:

$$S(f_n) = A(f_n) \,.\, d^{-\frac{1}{2}} \,.\, \exp\,(-\alpha_f \,.\, d) \tag{12}$$

$$\left.\begin{aligned} &= E(f_n) \,.\, \ddot{\phi}(f_n)^{\frac{1}{2}} \qquad (13)\\ &= E(f_n) \,.\, (d/c)^{\frac{1}{2}} \,.\, f_c(f^2 - f_c{}^2)^{-3/4} \qquad (14)\end{aligned}\right\} \quad \text{for } f_n \neq f_T$$

The evaluation of the form of this expression requires an estimate of the cut-off frequency of the waveguide $f_c = c/2h$, which can only be obtained from the

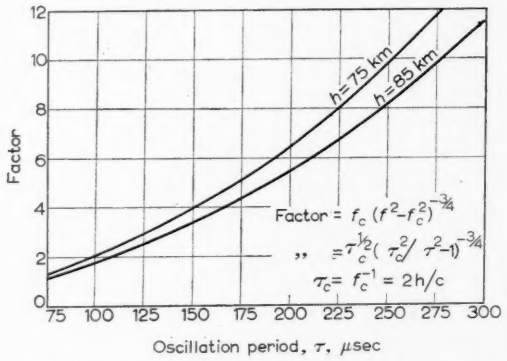


Fig. 6. Frequency dependence of the phase variation factor due to propagation.

waveform analysis in the event of an independent and reliable estimate of the distance of the discharge such as may be available as an individual waveform fix by a Sferics network. Preliminary analyses of day-time waveforms for which such fixes are available suggest it may be necessary to postulate an ionospheric layer at an effective height greater than 80 km. Spectra for the waveforms discussed earlier are shown in Fig. 7 deduced for an ionosphere height of 85 km and neglecting the $\ddot{\theta}(f_n)$ term for the waveforms showing irregularities.

The existence of a regular waveform is evidence of a simple original pulse. It is reasonable to conclude that the Fourier analyses of pulses giving such waveforms will be very similar in form, if not also in magnitude, and measurements on this one variety of waveform will be much more closely grouped than on waveforms in general. This average simple source spectrum may be deduced from the variation with distance of the oscillation amplitude $E(f)$ normalized to the value $E(F)$ at some reference frequency F, according to the formula:

$$\ln E(f)/E(F) = -(\alpha_f - \alpha_F) \,.\, d + \ln \{A(f) \,.\, \ddot{\phi}(F)^{\frac{1}{2}}/A(F) \,.\, \ddot{\phi}(f)^{\frac{1}{2}}\} \tag{15}$$

The relative attenuation coefficient $(\alpha_f - \alpha_F)$ with respect to the reference frequency is obtained directly as the slope of the plot and the normalized source spectrum $A(f)/A(F)$ from the variation of the intercept with frequency.

9. Conclusion

Smooth type waveforms may be classified upon inspection into regular and irregular forms. Accurate waveguide or pulse reflection analysis of the apparently regular waveforms reveals a subdivision of this group, the pseudo regular waveforms. These are characterized by a discontinuity in the analytical plots due to slight constant displacement of the latter group of points. Analyses of the irregular waveforms also exhibit discontinuities but the groups of analytical points do not lie on displaced portions of a single simple curve.

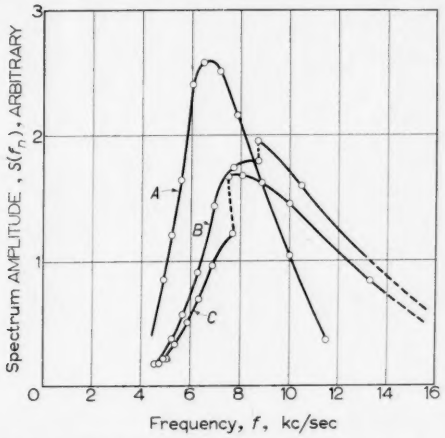


Fig. 7. Signal spectra of the respective waveforms.

The truly regular waveforms may be interpreted in terms of waveguide or pulse reflection aspects of the propagation of a particularly simple pulse, provided the erroneous assumption that crests of the waveform must satisfy the reflection theory formula is discarded. Pseudo regular waveforms can be satisfactorily explained from either viewpoint, when more complicated source conditions are considered and it is not assumed that the discharge in the lightning channel is effectively instantaneous. Consideration of the waveforms to be expected from discharges having a pulsating current discloses a probable explanation of the irregular waveforms.

Amplitude analysis of the waveforms provides a very simple method of obtaining the individual waveform spectra, which may be averaged to obtain propagation attenuation coefficients and the mean spectrum at the source.

The conclusions of Caton and Pierce diverge from those of other workers, who are satisfied by interpretation on one or other of the theories which are seen to be mutually equivalent. This is possibly because of greater stringency in deciding an acceptable fit with theoretical predictions together with unfortunate discouragements owing to the absence of early order crests and disagreement with Sferics information. For waveforms originating at such great distances, the time relations for the early order reflections require amendment because the first-, second- and third-order rays become tangential to the earth for propagation distances of approximately 2000, 4000 and 6000 km respectively. More relevant

VOL. 14 1959

in the present context however is the fact that at distances near and beyond these critical limits the respective rays must propagate partially as ground rays with correspondingly high attenuation so that it is not surprising to find the early orders missing from these waveforms of very distant origin. The Sferics network uses a system of triangulation with a baseline restricted by the limits of convenient sites and telephonic interconnection to about 500 km. Because of experimental error in direction finding the zone of reasonably accurate fixing is restricted to 2500 km. For storms considerably beyond this range the storm distance estimated from Sferics data is of little significance and agreement of estimates is not to be expected. Unfortunately specific data whereby this hypothesis might be justified was not given by the authors. In view of the indeterminacy of the true reflection order for waveforms of short duration (HORNER and CLARKE, 1955) CATON and PIERCE's invalid proviso of peaks corresponding to sky pulses would seem to be irrelevant. The importance of their work in the present context lies in the attention drawn to those smooth waveforms which do not conform to simple pulse reflection theory and which was amongst the provocations to the present investigation.

Acknowledgements—The author is grateful to Professor L. F. BATES for the provision of research facilities in the Physics Department at the University of Nottingham, where the above study forms part of a continuing programme of atmospheric investigation.

REFERENCES

APPLETON E. V. and CHAPMAN F. W.	1937	*Proc. Roy. Soc.* **A 158,** 1.
BUDDEN K. G.	1951	*Phil. Mag.* **42,** 1.
CATON P. G. F. and PIERCE E. T.	1952	*Phil. Mag.* **43,** 393.
HALES A. L.	1948	*Proc. Roy. Soc.* **A 193,** 60.
HEPBURN F.	1957	*J. Atmosph. Terr. Phys.* **10,** 121.
HEPBURN F.	1958	*J. Atmosph. Terr. Phys.* **12,** 1.
HORNER F. and CLARKE C.	1955	*J. Atmosph. Terr. Phys.* **7,** 1.
LABY T. H., MCNEILL J. J., NICHOLLS F. G. and NICKSON A. F. B.	1940	*Proc. Roy. Soc.* **A 174,** 145.
LABY T. H., NICHOLLS F. G., NICKSON A. F. B. and MCNEILL J. J.	1938	*Nature, Lond.* **142,** 353.
LIEBERMAN L.	1956	*J. Appl. Phys.* **27,** 12, 1473.
LUTKIN F.	1939	*Proc. Roy. Soc.* **A 171,** 285.
MALAN D. J. and COLLENS H.	1937	*Proc. Roy. Soc.* **A 162,** 175.
SCHONLAND B. F. J., ELDER J. S., HODGES D. B., PHILIPS W. E. and VAN WYK J. W.	1940	*Proc. Roy. Soc.* **A 176.** 180.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 273 to 286. Pergamon Press Ltd. Printed in Northern Ireland

Some further analyses of *E*-layer measurements in South Africa during the solar eclipse of 25 December 1954

M. W. McElhinny
Department of Physics, University College of Rhodesia and Nyasaland,
Salisbury, Southern Rhodesia

(*Received* 12 *August* 1958)

Abstract—The *E*-layer measurements from three stations in South Africa during the solar eclipse of 25 December 1954 have been analysed from three points of view: (a) all the radiation comes from the visible disc, in which case $\alpha'_E = 1{\cdot}5 \times 10^{-8}$ cm^3 sec^{-1}; (b) 15 per cent of the radiation comes from outside the disc and $\alpha'_E = 4 \times 10^{-8}$ cm^3 sec^{-1}; and (c) the effect of two species of ion with differing recombination coefficients. It is shown that to neglect the effect of a more rapidly recombining species of ion may lead to an overestimate of the effect of the corona or of the limb-brightening.

VOL.
14
1959

1. Introduction

Vertical incidence ionospheric measurements were taken at three stations in South Africa during the annular solar eclipse of 25 December 1954. Measurements at Grahamstown, which was situated in the belt of annularity were made by the Department of Physics, Rhodes University and some preliminary results have already been published (Szendrei and McElhinny, 1956a,b). Readings at Cape Town and Johannesburg were made by the Telecommunications Research Laboratory of the South African Council for Scientific and Industrial Research. The circumstances of the eclipse at these three places for a height of 100 km are shown in Table 1, where $f_{\min}$ represents the unobscured fraction of the visible disc at the maximum phase.

At Grahamstown, readings were taken over thirty control days centring on the eclipse day every 15 min from 0500 to 1000 SAST and hourly intervals for the rest of the day. For the duration of the eclipse, readings were taken at $2\frac{1}{2}$ min intervals. Critical frequencies could be read to an accuracy of 0·01 Mc/s, but on the eclipse day on a few occasions could only be read to the nearest 0·02 Mc/s.

For Johannesburg and Cape Town, control data were taken from the "Monthly Bulletin of Ionospheric Characteristics at Johannesburg and Cape Town for December 1954," published by the Telecommunications Research Laboratory. Critical frequencies were only given to the nearest 0·1 Mc/s and hourly values only were available. Readings covering the 12 day period from 19 to 31 December were used as control data. On the eclipse day, readings were taken every minute and critical frequencies were read to the nearest 0·01 Mc/s, but on a few occasions could only be read to the nearest 0·05 Mc/s.

2. Interpretation of Results

From the critical frequency, f_0, the maximum electron density N, can be calculated from:

$$N = 1{\cdot}24 \times 10^4 f_0^2 \qquad (1)$$

where f_0 is expressed in Mc/s.

The variation of the maximum electron density is expressed by the usual ionospheric continuity equation:

$$\frac{dN}{dt} = q - \alpha' N^2 \tag{2}$$

where q is the rate of ion production per cm^3, and α' is the effective recombination coefficient. During an eclipse q must be replaced by fq where f is the residual fraction of the solar radiation.

Table 1. Circumstances of eclipse of 25 December 1954 at 100 km height

Station	Begins SAST	Middle	Ends SAST	f_{min}
Grahamstown Lat: 33°19′ S Long: 26°31′ E	0646	0800 to 0807 (annular phase)	0935	0·142
Johannesburg Lat: 26°10′ S Long: 28°05′ E	0639	0758	0929	0·320
Cape Town Lat: 34°9′ S Long: 18°19′ E	0647	0758	0922	0·204

SAST = South African Standard Time = UT + 2 hr.

It is usual as a first approximation to consider the sun to be a uniformly radiating disc and so f will represent the uncovered fraction of the visible disc. By selecting an appropriate value of α', q can be calculated from the control data and a theoretical curve plotted for the eclipse variation of N by following the gradient dN/dt from stage to stage during the eclipse. The resulting theoretical curves are rarely satisfactory. The time lag δt between the maximum obscuration and the occurrence of the minimum electron density N_m is usually small indicating a high value of α'. In addition N_m is higher than would be expected indicating a small value of α'. To explain this so-called N_m–δt anomaly there are three factors which might deviate from simple theory during an eclipse.

In the first place the obscuration factor f may not be equal to the unobscured fraction of the visible disc and more radiation may be concentrated in active areas on the solar disc. Secondly the effective recombination coefficient may vary during the course of the eclipse. Thirdly, some radiation might come from outside the visible disc.

3. The Distribution of Solar Ionizing Radiation

In a previous paper (Szendrei and McElhinny, 1956b) an attempt was made to deduce the distribution of active areas on the solar disc from the measurements

at Grahamstown. It was found necessary to postulate the existence of three sources, one on the west limb, one on the east limb, and an area situated in the north-western part of the disc. The value of α' chosen was 2×10^{-8} cm^3 sec^{-1}, and as was stated other values of α' do not materially affect this conclusion. The intensity of the sources will vary with different values of α', and their positions

Table 2. Grahamstown: variation of intensity of bright areas (percentage of total radiation) with α' (cm^3 sec^{-1})

	$1{\cdot}0 \times 10^{-8}$	$1{\cdot}5 \times 10^{-8}$	2×10^{-8}	4×10^{-8}	10^{-7}
West limb	12	8	5	4	3
East limb	3	6	8	8	9
Spot	20	16	14	10	8
Disc	65	70	73	78	80

only slightly. It was not possible to say whether the sources were in the northern or southern hemisphere, and they were consequently all shown as being in the northern hemisphere. Due to an error, the position of the solitary sunspot observed on 25 December, was plotted in the wrong place. The data from which its

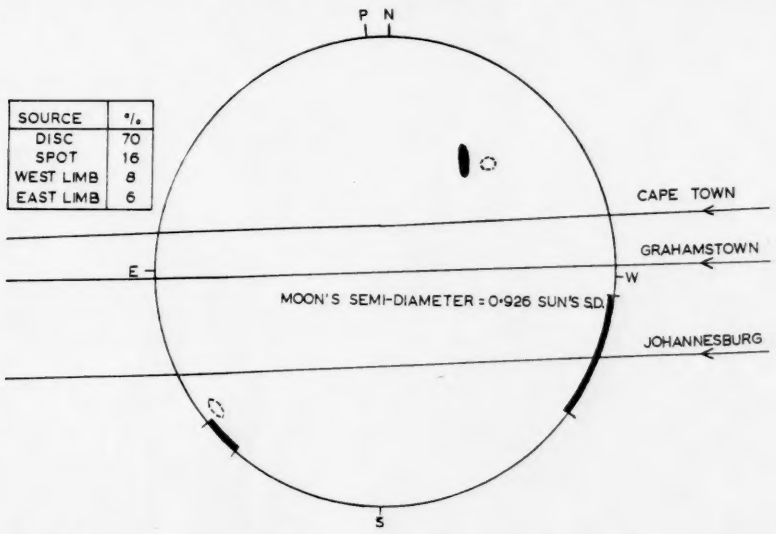


Fig. 1. Solar model—All radiation from visible disc. ---- Observed sunspots associated with enhancement in 5303 Å.

position was taken were plotted on heliographic co-ordinates, and since the pole of the sun was displaced about 6° east of the north point, its position will approximate more closely to that deduced from the ionospheric measurements.

Table 2 shows how the intensity of the three sources will vary with variation in the value of α' at Grahamstown.

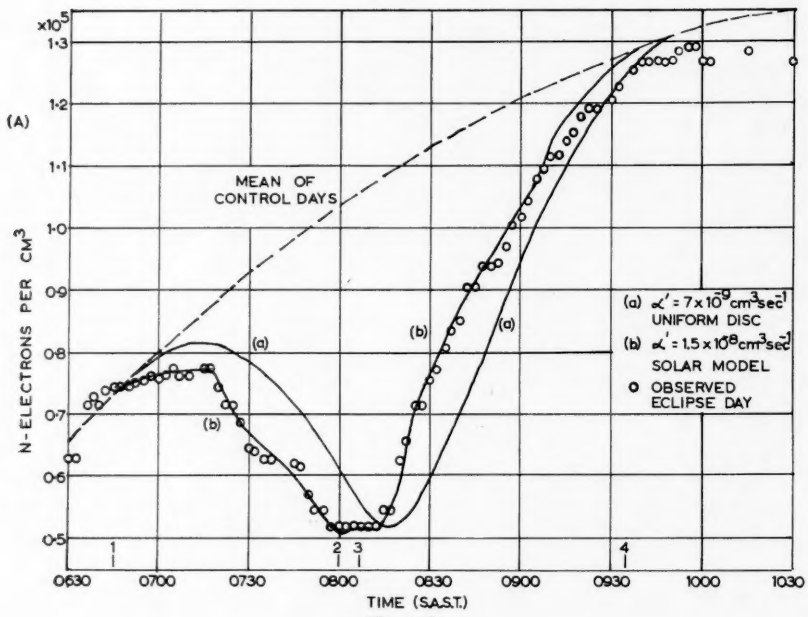

(A)
MEAN OF CONTROL DAYS
(a)
(b)
(a) α' = 7 x 10⁻⁹ cm³ sec⁻¹ UNIFORM DISC
(b) α' = 1.5 x 10⁻⁸ cm³ sec⁻¹ SOLAR MODEL
OBSERVED ECLIPSE DAY
N-ELECTRONS PER CM³
TIME (S.A.S.T.)

Fig. 2(A)

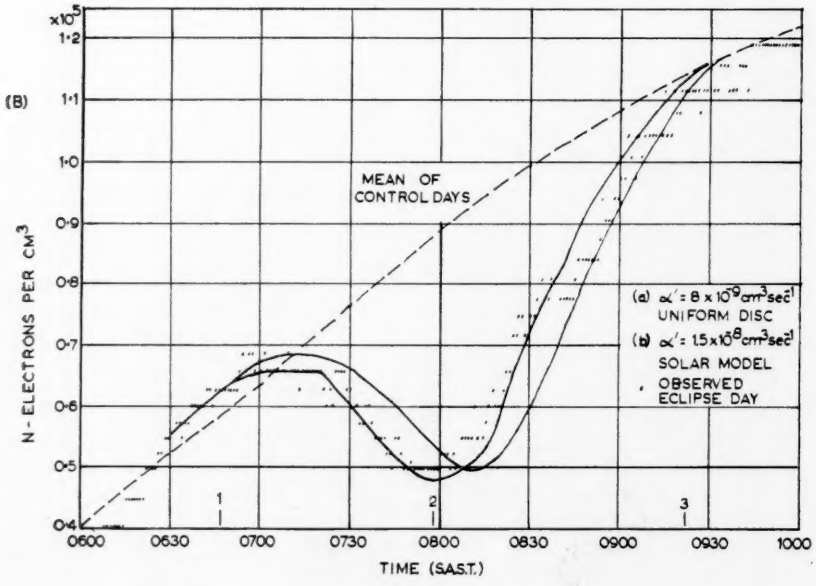

(B)
MEAN OF CONTROL DAYS
(a) α' = 8 x 10⁻⁹ cm³ sec⁻¹ UNIFORM DISC
(b) α' = 1.5 x 10⁻⁸ cm³ sec⁻¹ SOLAR MODEL
OBSERVED ECLIPSE DAY
N-ELECTRONS PER CM³
TIME (S.A.S.T.)

Fig. 2(B)

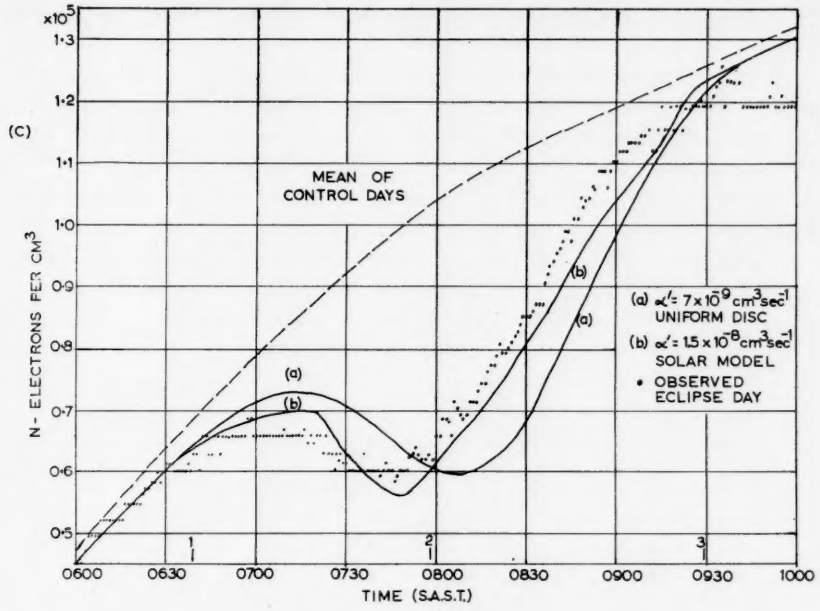


Fig. 2(C)

Fig. 2. Theoretical curves for solar model in Fig. 1. (A) Grahamstown; (B) Cape Town; (C) Johannesburg.

It is not possible to apply the method of MINNIS (1956a) to the three stations to deduce a solar model that would be the same for all three stations. However each of the five models shown in Table 2 were applied to the readings from Johannesburg and Cape Town to see if any would agree. There is no reason to suppose that the value of α' would be the same at all three stations, although one would not expect them to differ appreciably. It was found that the model corresponding to $\alpha' = 1{\cdot}5 \times 10^{-8}$ cm^3 sec^{-1} at Grahamstown gave the best fit at Johannesburg and Cape Town with the same value of α' at each place. It is necessary to place the east and west limb sources in the southern hemisphere whilst the spot remains in the northern hemisphere. The solar model is shown in Fig. 1. The value of α' can only at the best be given as $\alpha' = 1{\cdot}5 \pm 0{\cdot}3 \times 10^{-8}$ cm^3 sec^{-1}. Fig. 2 shows the theoretical curves for the three stations based on this model. The curves for a uniformly radiating disc are also included, and using the value of N_m as the criterion for determining α', the values agree very well with $\alpha' = 7 \times 10^{-9}$ cm^3 sec^{-1} at Grahamstown and Johannesburg and $\alpha' = 8 \times 10^{-9}$ cm^3 sec^{-1} at Cape Town. At Johannesburg and Cape Town it was found necessary to adjust the values of q before and after the eclipse to account for the departure from the mean of the control days. This is probably not a real effect, but due to the inadequate control data, available at only hourly intervals and to the nearest 0·1 Mc/s. The adjustments made lie within the standard deviation.

Since the publication of the previous paper (SZENDREI and MCELHINNY, 1956b) data has become available on the limb intensities of the Coronal Green line

(Trotter and Roberts, 1955) about 25 December 1954. A correlation has been found between the enhancement of the Coronal Green line (5303 Å) and the distribution of more intense regions of ionizing radiation on the disc as deduced from ionospheric measurements for example by Waldmeier (1947) and Minnis (1955). The coronal contour diagrams show that no radiation in 5303 Å was observed on either limb on 25 December. Radiation in 6374 Å was observed on both limbs but was only faint. The solitary spot was associated with an area of moderate intensity in 5303 Å, and it is possible to suppose that there was a very faint enhancement in 5303 Å just inside the east limb as shown in Fig. 1, close to the east limb source postulated. This area showed very bright intensity in 5303 Å when observed on the west limb 12 days later, but was not observed on the limb the previous 2 days. There seems no evidence therefore for the existence of the source on the west limb and very little for the east limb source.

4. Variation of the Recombination Coefficient

Some previous workers (e.g. McLeish, 1948) have found that the effective recombination coefficient changed during the eclipse from a high value at the beginning of the eclipse to a low value at the maximum phase. This was accounted for on the recombination theory of Massey (1937) and Bates *et al.* (1939), where it was supposed that the ratio of negative ions to electrons decreased during the eclipse causing a change in α'. More recent work by Bates and Massey (1946, 1947) has shown that probably dissociative recombination is the main process of electron removal in the E-layer. It is well known that solar ultra-violet and X-radiation is sufficient to ionize all atmospheric constituents (Bates, 1956), and therefore the existence of several species of ion is beyond doubt. Bates (1951) has shown that the diurnal variation of electron density may be completely controlled by a more slowly recombining species, and give no indication of a more rapidly recombining species. During an eclipse the more rapidly recombining species will disappear first, showing a high value of α' at the beginning of the eclipse and a low value determined by the slower species during the maximum phase. This might therefore account at least partly for the N_m–δt anomaly. Bates and McDowell (1957) have considered this theoretically for the case of a total eclipse. Following Bates and McDowell therefore, the sun will be considered a uniformly radiating disc, and it will be supposed that there are two species of ion whose concentrations are x and y per cm^3 at any time. Let q_x and q_y be their rates of production and α_x and α_y their respective recombination coefficients. On control days therefore:

$$\frac{dx_0}{dt} = q_x - \alpha_x x_0 N \tag{3}$$

$$\frac{dy_0}{dt} = q_y - \alpha_y y_0 N \tag{4}$$

where $N = (x_0 + y_0)$

and thus

$$\frac{dN}{dt} = \frac{dx_0}{dt} + \frac{dy_0}{dt} = q_x + q_y - N(\alpha_x x_0 + \alpha_y y_0) \tag{5}$$

Three cases have been considered:

Case (a) $\alpha_x = \alpha_y$, and the layer is a single species layer.

(b) $\frac{x_0}{y_0} = 4, \quad \alpha_x = 5 \times 10^{-9}, \quad \alpha_y = 10^{-7}$

(c) $\frac{x_0}{y_0} = 1 \quad \alpha_x = 10^{-9}, \quad \alpha_y = 10^{-7}$

In the latter two cases, q_x and q_y can be calculated from the control values of N and dN/dt, the assumption being made that $dx_0/dt = dy_0/dt$. Since dN/dt is small, this does not materially effect the analysis. Using these values of q_x and q_y deduced from control data, the gradients dx/dt and dy/dt can be followed from stage to stage during the eclipse, and hence x, y and consequently N can be plotted on a theoretical curve using the equations for the eclipse.

$$\frac{dx}{dt} = fq_x - \alpha_x x N \tag{6}$$

$$\frac{dy}{dt} = fq_y - \alpha_y y N \tag{7}$$

VOL. 14 1959

The three cases are shown in Fig. 3 in which the theoretical curves so deduced are plotted with those observed for the results from Grahamstown. In case (a) $\alpha_x = \alpha_y = 7 \times 10^{-9}$ was found to give the same minimum N_m as that observed corresponding to the case of a single species layer.

In cases (b) and (c), the theoretical curves approach that observed more closely but they do not account for the rapid drop in N just after 0717 SAST. Only one active area on the sun was present on the eclipse day, associated with a solitary sunspot in the north-west part of the disc as shown in Fig. 1.

The best fit to the observed points was found by supposing that 80 per cent of the radiation comes from the visible disc and 20 per cent from the active area. This case is shown in Fig. 4. In case (b) the N_m–δt anomaly is not sufficiently accounted for, and in case (c) whereas the observed points are accounted for before and during annularity, the final recovery during the last stages of the eclipse is completely controlled by the more slowly recombining species $\alpha_x = 10^{-9}$ cm^3 sec^{-1}, and does not account for that observed.

The results show that it is unlikely that the N_m–δt anomaly in this eclipse can be completely accounted for by the presence of a more rapidly recombining species of ion, however it must be pointed out that neglect of this factor may lead to an overestimate of other factors such as the distribution of solar radiation on the solar disc.

5. Radiation from Outside the Visible Disc

One of the most remarkable features of all eclipse measurements (Ratcliffe, 1956) is that in practically all cases the minimum electron density, N_m, occurs during the maximum phase of the eclipse, but fails to reach the expected lowest value by about 10 or 15 per cent. If the explanation is attributed to the distribution of intense regions on the sun it would seem a remarkable coincidence, when

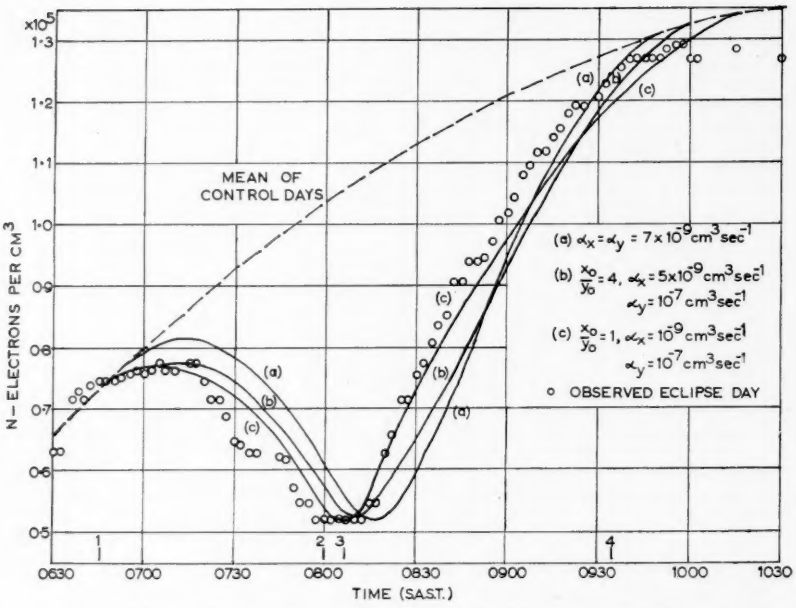


Fig. 3. Grahamstown—variation of recombination coefficient.

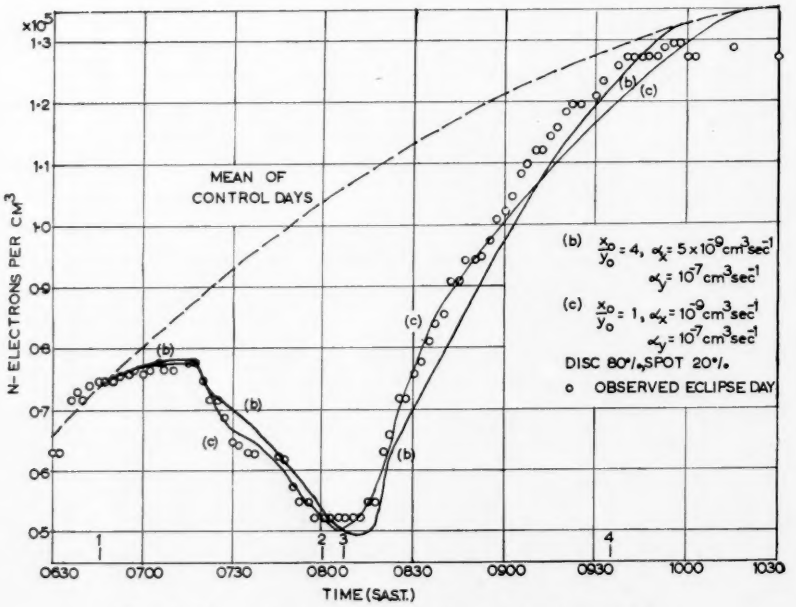


Fig. 4. Grahamstown—variation of recombination coefficient.

one would expect a chance distribution. It seems likely as suggested by PIDDINGTON (1951) and RATCLIFFE (1956) that some radiation must come from the corona. A solar model is proposed as shown in Fig. 5 in which 15 per cent of the total radiation comes from outside the visible disc, which has been extended by about 0·1 solar diameter at the equator in such a way that the intensity per unit area coming from outside the visible disc is the same as the disc itself. In addition some allowance must be made for the active area in the north-west quadrant.

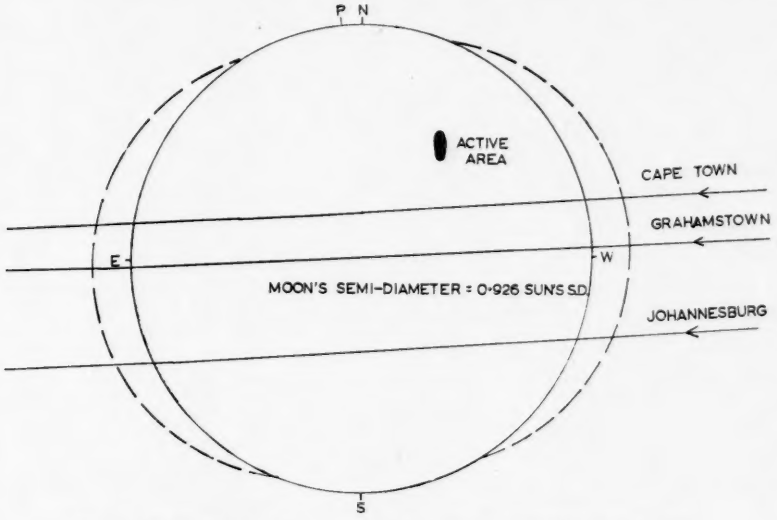


Fig. 5. Solar model—15 per cent radiation from outside visible disc.

This model was applied to the three stations as shown in Fig. 6. At Grahamstown the best fit is for $\alpha' = 4 \times 10^{-8}$ cm^3 sec^{-1} if from 5 to 10 per cent comes from the spot, whilst equally good agreement can be obtained with $\alpha' = 2 \times 10^{-8}$ cm^3 sec^{-1} and 15 per cent attributed to the spot.

At Johannesburg, if from 5 to 10 per cent comes from the spot then $\alpha' = 4 \times 10^{-8}$ cm^3 sec^{-1} and higher values of α' would not make much difference to the fit. However, at Cape Town if from 5 to 10 per cent comes from the spot $\alpha' = 4 \times 10^{-8}$ cm^3 sec^{-1}, and higher values of α' do not give such a good fit.

6. Discussion

MINNIS (1956a,b) has been able to account for the results from Khartoum and Ibadan during the 1952 eclipse, and those from Inverness and Slough during the 1954 eclipse very satisfactorily on the assumption that no radiation comes from outside the visible disc. In the former case α' was found to be $1{\cdot}2 \pm 0{\cdot}1 \times 10^{-8}$ cm^3 sec^{-1} when 36 per cent came from two active areas. BONNET, HUNAERTS and NICOLET (1957) in analysing amongst others the results from Khartoum and Ibadan, showed that if 15 per cent of the radiation came from outside the visible disc and only 11 per cent from two spots on a model similar to that proposed in this paper, $\alpha' = 4 \times 10^{-8}$ cm^3 sec^{-1}. The two results are similar to that found in this paper.

VOL. 14 1959

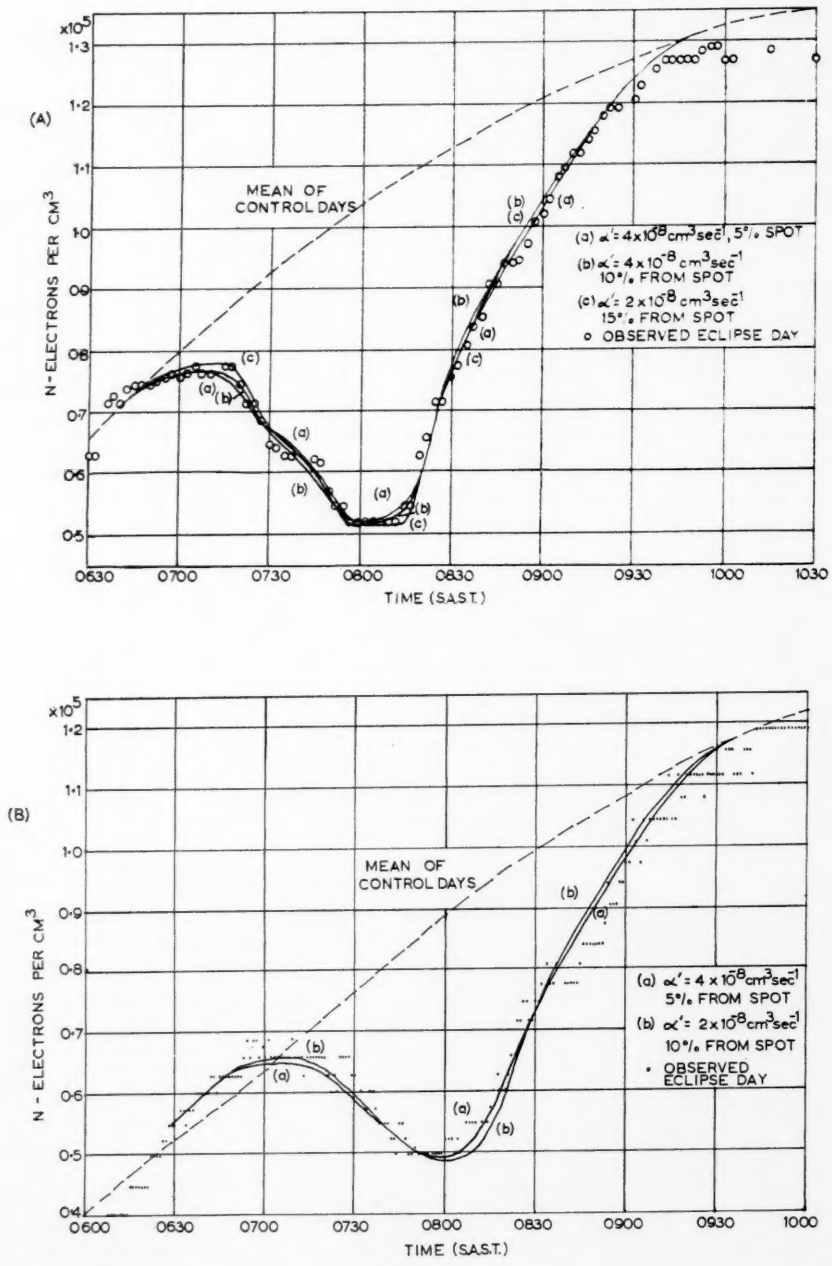

(A)
x10^5
1·3
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
N - ELECTRONS PER CM^3
MEAN OF CONTROL DAYS
(a) α' = 4x10^-8 cm^3 sec^-1, 5°/o SPOT
(b) α' = 4x10^-8 cm^3 sec^-1 10°/o FROM SPOT
(c) α' = 2x10^-8 cm^3 sec^-1 15°/o FROM SPOT
o OBSERVED ECLIPSE DAY
0630
0700
0730
0800
0830
0900
0930
1000
1030
TIME (S.A.S.T.)
(B)
x10^5
1·2
1·1
1·0
0·9
0·8
0·7
0·6
0·5
0·4
N - ELECTRONS PER CM^3
MEAN OF CONTROL DAYS
(a) α' = 4 x 10^-8 cm^3 sec^-1 5°/o FROM SPOT
(b) α' = 2x10^-8 cm^3 sec^-1 10°/o FROM SPOT
• OBSERVED ECLIPSE DAY
0600
0630
0700
0730
0800
0830
0900
0930
1000
TIME (S.A.S.T.)

Fig. 6

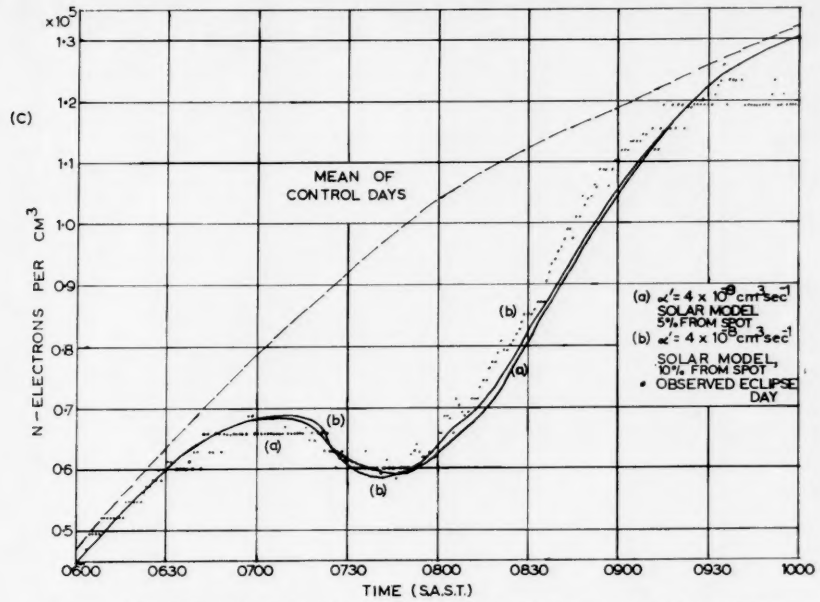

(C)
MEAN OF CONTROL DAYS
N-ELECTRONS PER CM3
TIME (S.A.S.T.)
(a) α'= 4 x 10^-8 cm^3 sec^-1 SOLAR MODEL 5% FROM SPOT
(b) α'= 4 x 10^-8 cm^3 sec^-1 SOLAR MODEL 10% FROM SPOT
OBSERVED ECLIPSE DAY

Fig. 6 (Continued). Theoretical curves for solar model in Fig. 5. (A) Grahamstown; (B) Cape Town; (C) Johannesburg.

VOL. 14 1959

Minnis (1956b) has pointed out that the absolute intensity of the uniform disc component is about the same for both the 1952 and 1954 eclipses. Using the notation of Minnis, the diurnal variation of N may be expressed as:

$$\frac{dN}{dt} = q_0 \cos^n \chi - \alpha' N^2 \tag{8}$$

Rearranging the terms:

$$J = \frac{q_0}{\alpha'} = \left(N^2 + \frac{1}{\alpha'} \frac{dN}{dt} \right) \sec^n \chi \tag{9}$$

where q_0 = rate of generation of electron per cm^3 for an overhead sun.
χ = solar zenith angle

For a Chapman layer $n = 1$, but it is usually found that $n \neq 1$, and the appropriate value of n is used as deduced from the diurnal variation on control days.

Using the noon values of N from the control days at the three stations when $dN/dt = 0$, values of J were deduced from equation (9) and are shown in Table 3 together with the values of n for each station.

Taking the mean value of J as 220×10^8 cm^{-6},

$$q_0 = 220\ \alpha' \times 10^8 \text{ cm}^{-3} \text{ sec}^{-1}$$

During the 1952 eclipse at Khartoum $q_0 = 258\alpha' \times 10^8$ cm^{-3} sec^{-1}.

Table 4 shows the Uniform Component of q_0 as deduced from the two eclipses on the two models that (A) all the radiation comes from the visible disc, and (B) 15 per cent comes from outside the visible disc.

The agreement for the values of B is better than that for A, but the result is not conclusive, since the value of α' and J for the 1954 eclipse are not as accurate as those of Minnis for the 1952 eclipse. It would be interesting to analyse the results from Slough and Inverness for the June 1954 eclipse in terms of a model with 15 per cent of the radiation coming from outside the visible disc.

Table 3. Values of n and J

Station	n	$J(10^8$ cm$^{-6})$
Grahamstown	1·18	211
Cape Town	1·12	220
Johannesburg	1·20	229

Table 4. Absolute intensity of uniform component of solar radiation. A: all radiation from visible disc. B: 15 per cent radiation from outside visible disc.

Eclipse	q_0 (cm^{-3} sec^{-1})	A (cm^{-3} sec^{-1})	B (cm^{-3} sec^{-1})
Feb. 1952	$258\alpha'$	$198(\alpha' = 1{\cdot}2 \times 10^{-8})$	$916(\alpha' = 4 \times 10^{-8})$
Dec. 1954	$220\alpha'$	$231(\alpha' = 1{\cdot}5 \times 10^{-8})$	$836(\alpha' = 4 \times 10^{-8})$

Minnis (1955, 1956a) has attempted to correlate the intensity of the west limb source of 25 per cent of the total radiation during the 1952 eclipse with changes in J several days following the eclipse. If this source was only of a temporary importance, there should be a drop in J following the eclipse day when this source disappeared behind the disc. The observed drop in J was only 13 per cent. Although it is possible that other sources may have in the meanwhile appeared on the east limb, it is interesting to note that the analysis of Bonnet, Hunaerts and Nicolet requires only 11 per cent of the radiation to be concentrated on the west limb, agreeing well with the observed drop in J.

According to Allen (1948), the intensity I_R of the E-layer ionizing radiation for a twelve-month mean sunspot number $\bar{R}$ is given by

$$I_R = I_0(1 + 0{\cdot}0097\bar{R}) \tag{10}$$

For December 1954 $\bar{R} = 10$ and hence $I_R = 1{\cdot}097I_0$. This indicates that the level of activity during December 1954 was only about 10 per cent above that corresponding to zero sunspot number, or what might be interpreted as the minimum level of solar activity. This does not seem consistent with the solar model with only 70 per cent of the radiation emitted uniformly from the disc, but agrees well with the model in which 15 per cent comes from outside the visible disc, when only 5 to 10 per cent need be attributed to a single active area. These factors seem to

point in favour of some radiation coming from outside the visible disc and consequently the higher value of α'.

One of the features of the records from Cape Town was the number of "critical frequencies" which were observed. On occasions up to six critical frequencies could be seen on the $h'f$ record, three of which were from the E-region. These cusp-like anomalies have been shown by MUNRO and HEISLER (1956, 1958) to be due to travelling ionospheric disturbances of a wave-like nature producing tilts in the ionospheric layers and causing echoes to be returned from directions other than the vertical. GLEDHILL (1958) has made a preliminary analysis of the measurements from Grahamstown and has shown that it is unlikely that reflections during the middle of the eclipse are returned from points overhead. Since they would be returned from points where the electron density is necessarily higher it seems likely that the apparent critical frequency will be higher than the true one. This would give the appearance of some radiation coming from outside the visible disc, but would indicate that the value of α' could be as high as 4×10^{-8} cm^3 sec^{-1}, and not support the lower values deduced by MINNIS.

7. CONCLUSION

The E-layer measurements at three stations in South Africa have been analysed from three points of view. If all the ionizing radiation comes from the visible disc then $\alpha'_E = 1{\cdot}5 \times 10^{-8}$ cm^3 sec^{-1} gives reasonable agreement with that observed with a model in which 70 per cent of the radiation comes uniformly from the disc. If 15 per cent comes from outside the disc then $\alpha'_E = 4 \times 10^{-8}$ cm^3 sec^{-1} and only from 5 to 10 per cent need be attributed to a solitary sunspot. It is very difficult to distinguish between the two cases. However there is little if any evidence for the existence of the two sources on the limb which have to be postulated if all the radiation comes from the visible disc, and this would support the idea that about 15 per cent of the radiation comes from outside the visible disc. It is pointed out that to neglect the presence of a more rapidly recombining species of ion may lead to an overestimate of the effect of the corona or of sources on the sun's limb. Finally, it seems possible that the presence of oblique echoes during the middle of the eclipse, even in apparently simple $h'f$ records, may give the appearance of radiation coming from outside the disc and consequently support the higher value of $\alpha_E' = 4 \times 10^{-8}$ cm^3 sec^{-1}.

Acknowledgements—The author wishes to thank Professor J. A. GLEDHILL of Rhodes University for making available the measurements from Johannesburg and Cape Town, and Mr. C. M. MINNIS of the DSIR Radio Research Station, Slough, for forwarding the coronal limb intensities.

REFERENCES

ALLEN C. W.	1948	*Terr. Magn. Atmos. Elect.* **53,** 433.
BATES D. R.	1951	*Proc. Phys. Soc. Lond.* B **64,** 805.
BATES D. R.	1956	*Solar Eclipses and the Ionosphere* p. 184 Pergamon Press, London.
BATES D. R., BUCKINGHAM R. A., MASSEY H. S. W. and UNWIN J. J.	1939	*Proc. Roy. Soc.* A **170,** 322.

Bates D. R. and Massey H. S. W.	1946	*Proc. Roy. Soc.* A **187,** 261.
Bates D. R. and Massey H. S. W.	1947	*Proc. Roy. Soc.* A **192,** 1.
Bates D. R. and McDowell M. R. C.	1957	*J. Atmosph. Terr. Phys.* **10,** 96.
Bonnet G., Hunaerts J. and Nicolet M.	1957	*Bull. Acad. Roy. Sci. Col.* (N.S.) **3,** 964.
Gledhill J. A.	1958	*Conf. S. Afr. Inst. Phys. Pretoria, July* 23–25.
Massey H. S. W.	1937	*Proc. Roy. Soc.* A **163,** 542.
McLeish C. W.	1948	*Canad. J. Res.* A **26,** 137.
Minnis C. M.	1955	*J. Atmosph. Terr. Phys.* **6,** 91.
Minnis C. M.	1956a	*J. Atmosph. Terr. Phys.* **9,** 36.
Minnis C. M.	1956b	*J. Atmosph. Terr. Phys.* **9,** 201.
Munro G. H. and Heisler L. H.	1956	*Aust. J. Phys.* **9,** 343.
Munro G. H. and Heisler L. H.	1958	*J. Atmosph. Terr. Phys.* **12,** 57.
Piddington J. H.	1951	*J. Geophys. Res.* **56,** 409.
Ratcliffe J. A.	1956	*Solar Eclipses and the Ionosphere* p.1. Pergamon Press, London.
Szendrei M. E. and McElhinny M. W.	1956a	*J. Atmosph. Terr. Phys.* **8,** 108.
Szendrei M. E. and McElhinny M. W.	1956b	*Solar Eclipses and the Ionosphere* p. 74. Pergamon Press, London.
Trotter D. E. and Roberts W. O.	1955	*Report No.* HAO-26. High Altitude Observatory, Colorado.
Waldmeier M.	1947	*Terr. Magn. Atmos. Elect.* **52,** 333.

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 287 to 295. Pergamon Press Ltd. Printed in Northern Ireland

Geomagnetic activity following large solar flares

CONSTANCE S. WARWICK* and R. T. HANSEN†

(*Received* 6 *September* 1958)

Abstract—Superposed epoch analysis of geomagnetic index around day of occurrence of large central flares shows a significant increase in geomagnetic activity following flares selected according to: (a) area of the flare and (b) "importance" of the flare. The association does not extend to flares of lesser importance, 2. During the period 1950–1954, near sunspot minimum and roughly the same period studied by WATSON, there was no tendency for geomagnetic disturbance to follow within 2–3 days after the fourteen reported class 3 and 3+ flares.

1. INTRODUCTION

A STUDY by WATSON (1957) of geomagnetic activity following geomagnetic solar flare effect (s.f.e.) or crochet confirms the results of an earlier study by VAN SABBEN (1952) and suggests that there is no significant preferred tendency for geomagnetic disturbance to follow solar flares. Earlier work by NEWTON (1943), on the other hand, demonstrated a pronounced increase in geomagnetic activity about 2 days after major flares, identified as "importance 3+". ELLISON, in a recent note (1958), has re-emphasized NEWTON's result, and pointed out that WATSON, in selecting cases of s.f.e. has not followed NEWTON's criterion of major centrally located flares. Other investigations (SIMON, 1956; DODSON and HEDEMAN, 1958) have demonstrated the significance of radio-noise association in the flare–geomagnetic storm relation. We wish to report here on several studies of the geomagnetic effect of solar flares, identified solely on the basis of optical observations.

VOL. 14 1959

The superposed-epoch method of analysis was employed to show the run of the mean value of the geomagnetic character figure, A_p, from 14 days before to 14 days after the flare day. The effect of autocorrelation of A_p for adjoining days and days 27 days apart was taken into consideration in determining the statistical significance of the maxima in mean A_p and in frequency of $A_p \geq 30$ (BELL and GLAZER, 1957). A modified t-test was used to estimate the upper limit of the probability that a value as high or higher than the observed maximum would occur by chance for one or more of the twenty-nine values, by treating them as independent. We believe that these values are conservative, especially for mean A_p.

2. FLARES SELECTED ACCORDING TO AREA

For the first study, zero days were defined by the occurrence of flares with area greater than six-hundred millionths of the solar hemisphere, reported by the High Altitude Observatory (13 February 1950 to 13 February 1951) or by the Sacramento Peak Observatory (28 February 1951 to 11 April 1958). The list was limited to flares within 45° of the central meridian and areas were not corrected for foreshortening. No other criterion, other than flare areas, measured from $H\alpha$ filtergrams, entered into the selection. In Fig. 1, we see that there was a

* National Bureau of Standards, Boulder, Colorado, U.S.A.
† High Altitude Observatory, Boulder, Colorado, U.S.A.

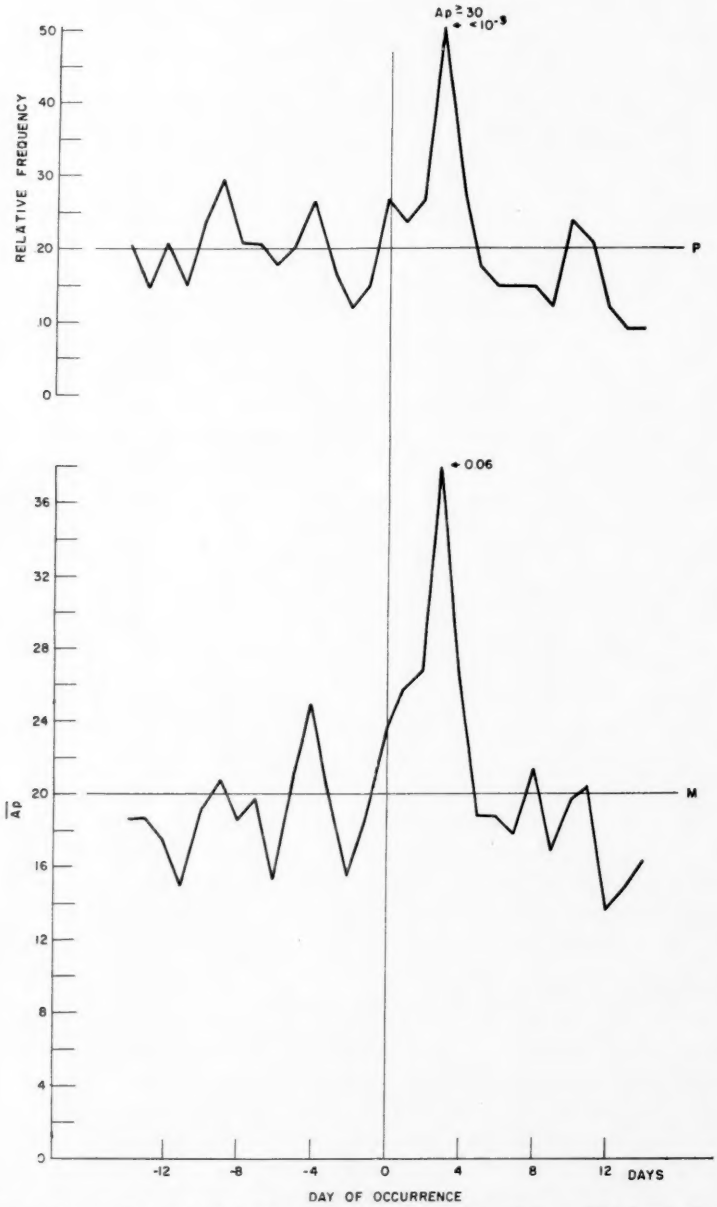


Fig. 1. Mean value of A_p (*below*) and frequency of values of $A_p > 30$ (*above*), about day of occurrence of flare of area 600×10^{-6}, thirty-four cases, 1950–1957.

substantial increase in the mean value of the magnetic index on the third day following these thirty-four flares. In Fig. 1 is also plotted the "frequency of occurrence" of A_p values ≥ 30 (roughly the lower limit of a definite magnetic storm) for each of the 14 days before and after the flare. Again, on the third day after the flare, there is an increase in the frequency of daily magnetic indices ≥ 30.

3. Flares Selected According to Optical "Importance"

For the second part of our study, the flare data were taken from the IAU Quarterly Bulletin of Solar Activity (1940–June 1956) and the NBS-CRPL F-Series Bulletins, Part B, (July 1956–October 1957). Because reports of flare importance by the various observatories do not always agree, we arbitrarily accepted the greatest importance reported by *any* station. We thus compiled a list of "Importance 3" flares within 45° of the central meridian. Days with more than one such flare entered the analysis twice.

The list was divided into three time periods: Jan. 1940–Feb. 1950, Mar. 1950–Dec. 1954, and Jan. 1955–Oct. 1957. There were no entries (that is, no "importance 3" flares) during the years 1953 and 1954. The superposed epoch plots for each period are seen in Figs. 2, 3 and 4, respectively. For the first and third periods, we see again high values both of mean A_p, and of frequency of $A_p \geq 30$, and of $A_p \geq 80$ on the second or third days after the flare. The unexpected maximum at day $+12$ for the second period (Fig. 3) seems to be caused by a coincidence between "importance 3" flares and geomagnetic disturbance, each event showing strong 27-day recurrence. Eight of the fourteen cases fit into this recurrence pattern. The period 1950–1954 falls near minimum of the solar activity cycle, when recurrent M-region storms dominate geomagnetic activity. The geomagnetic effect following flares selected by area in the period 1950–1957 (Fig. 1) is more significant and larger than that found for flares selected by "importance" during the same period (Figs. 3 and 4).

The period Jan. 1940–21 Feb. 1950, covered by the study shown in Fig. 2, coincides with the period for which Newton (1943) and Newton and Jackson (1951) published lists of intensity 3 and 3+ flares. For comparison with our study, we present in Fig. 5 the values of mean A_p, and of frequency of $A_p \geq 30$ and of $A_p \geq 80$, for central flares from these lists. We note here that Newton's and Newton and Jackson's lists represent a critical selection from among the various optical observations, so that many of the flares on our list do not appear on theirs. They also took into consideration other aspects of the flare, such as geomagnetic crochet, radio communication fadeout, and radio burst. Furthermore, some of the events on Newton and Jackson's list (1951) of 3 and 3+ flares were not observed optically, but were deduced from the flare-associated terrestrial effects. Such events do not appear on our list. Three of the twenty-two flares given by Newton or Newton and Jackson were not on our list; sixty of the seventy-nine flares in our study were not on their lists.

4. Importance "2" Flares

Next we took from the IAU Quarterly Bulletins the dates of all central flares reported as "importance 2", in the interval Jan. 1940–July 1956. The greatest

VOL. 14 1959

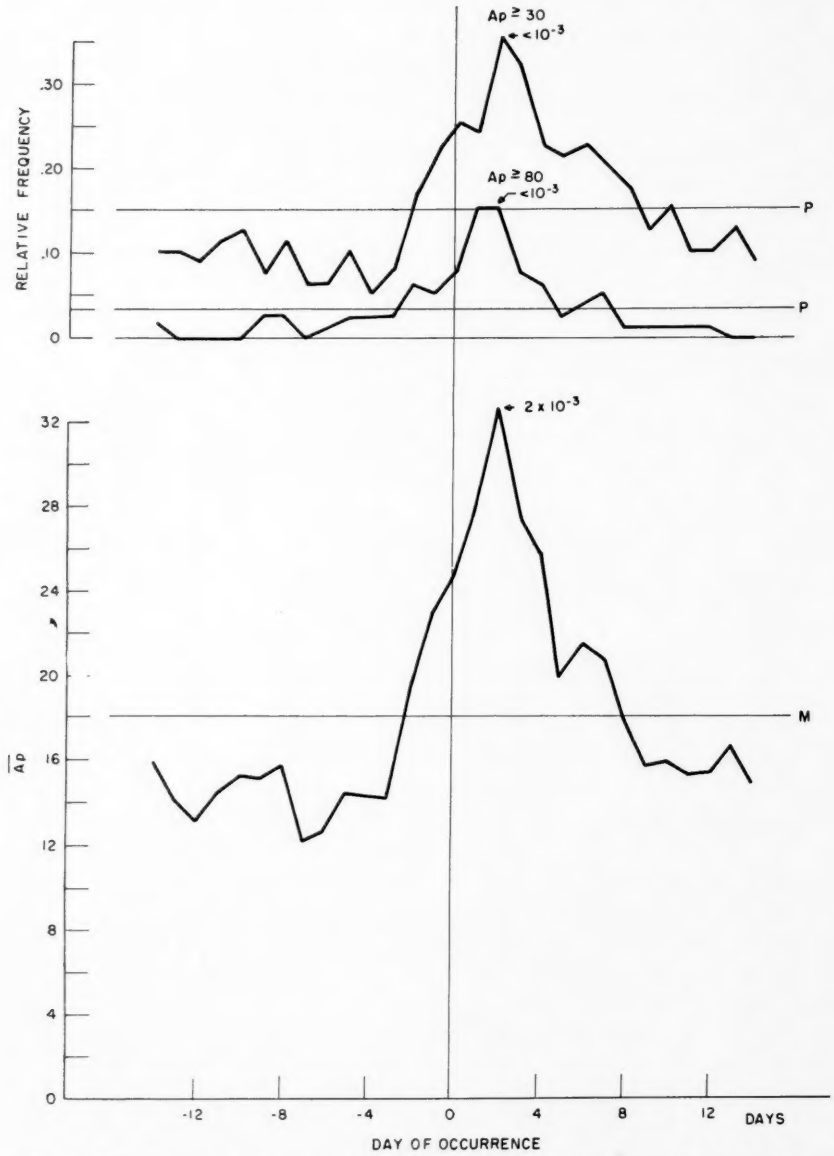


Fig. 2. Mean value of A_p (*below*) and frequency of values of $A_p > 30$ and of $A_p > 80$ (*above*) about day of occurrence of central flare of "importance 3" or "3+". Seventy-nine cases 1940–1950.

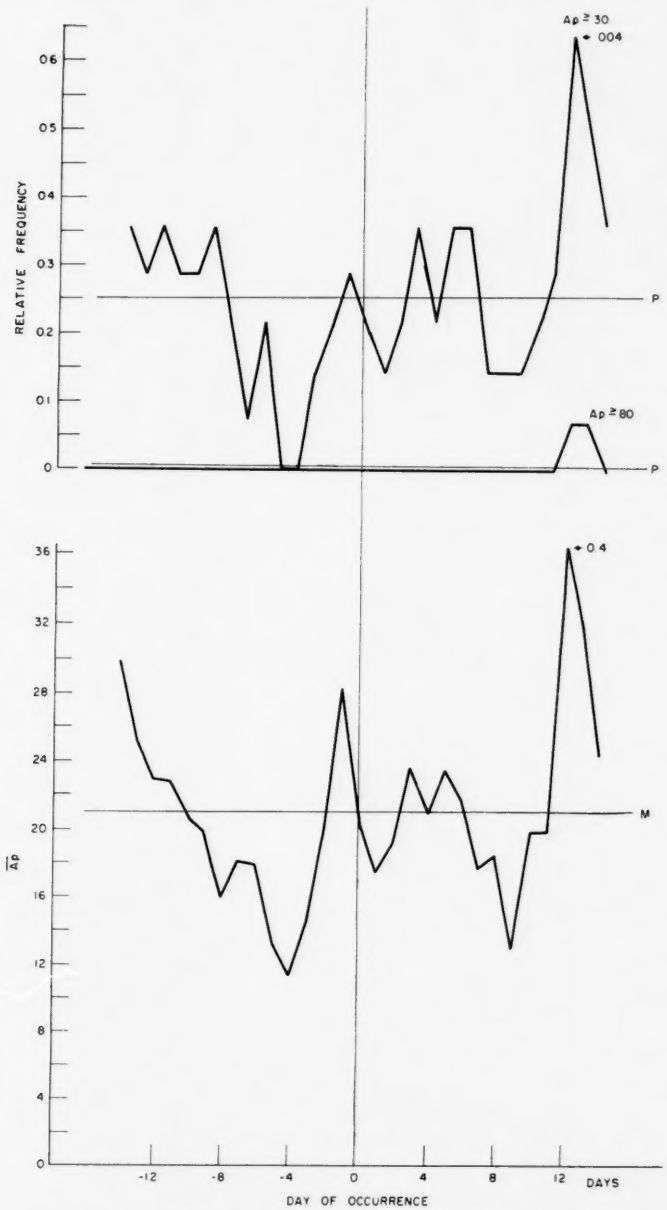


Fig. 3. Mean value of A_p (*below*) and frequency of values of $A_p > 30$ and of $A_p > 80$ (*above*) about day of occurrence of central flare of "importance 3" or "3+". Fourteen cases, 1950–1952.

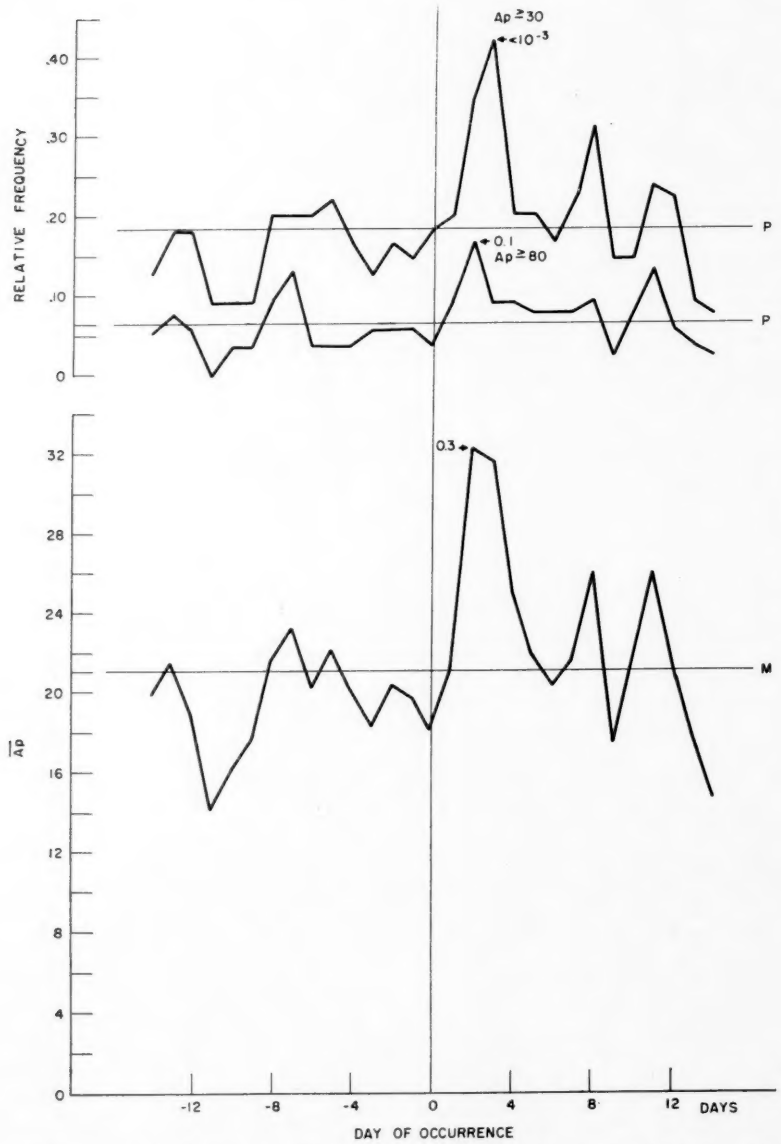


Fig. 4. Mean value of A_p (*below*) and frequency of values of $A_p > 30$ and of $A_p > 80$ (*above*) about day of occurrence of central flare of "importance 3" or "3+". Fifty-five cases, 1955–1957.

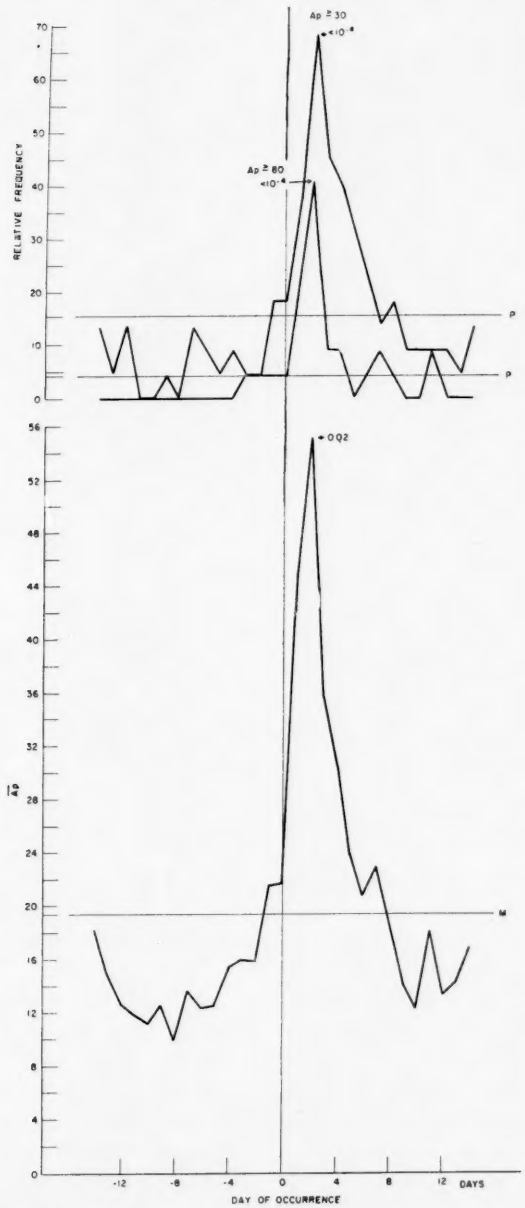


Fig. 5. Mean value of A_p (*left*) and frequency of values of $A_p > 30$ and of $A_p > 80$ (*right*) about day of occurrence of central flare of "importance 3" or "3+" selected by NEWTON. Twenty-two cases, 1940–1950.

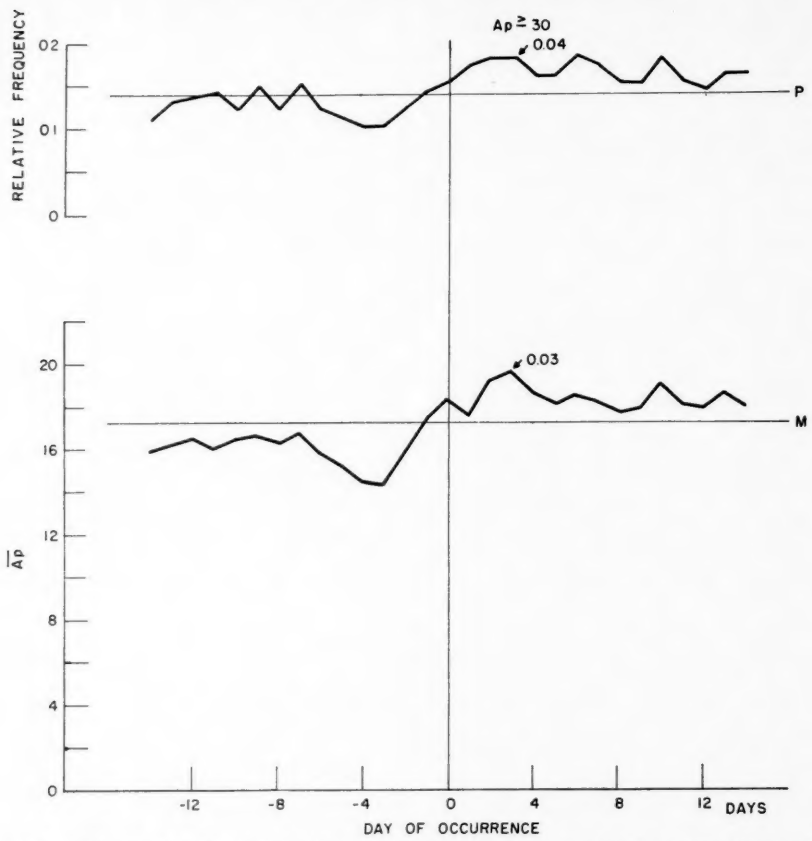


Fig. 6. Mean value of A_p (*below*) and of frequency of $A_p > 30$ (*above*) about day of occurrence of central flare of "importance 2" (no flare of "importance 3" or "3+" on same day). Six-hundred and twenty cases, 1942–1956.

importance given by *any* observatory was used and the date was excluded if an "importance 3" flare was listed for the same date. In Fig. 6, we see that the association between geomagnetic storminess and solar flares is largely lost for the 620 flares of lesser importance.

5. Conclusions

From these brief studies we conclude that:

(1) There is a definite tendency for geomagnetic disturbance to follow the largest solar flares, at least near the time of maximum of the sunspot cycle.

(2) The flare area (in our study uncorrected for fore-shortening), is probably a more reliable index of the likelihood of subsequent geomagnetic disturbance than the flare "importance".

(3) The period 1950–1954, nearly the same period as studied by WATSON (1951), is remarkable in not showing geomagnetic disturbance following 2 or 3 days after reported class 3 and 3+ flares. It seems possible that WATSON's negative result is due to his choice of years of low solar activity for his study. It is also possible that

VOL
14
195

during periods of relatively little solar activity, observers have tended to overestimate the optical "importance" of the larger flares.

References

Bell B. and Glazer H.	1957	*Smithson. Contr. Astrophys.* **2,** 51.
Dodson H. and Hedeman E. R.	1958	*J. Geophys. Res.* **63,** 77.
Ellison M. A.	1958	*J. Geophys. Res.* **12,** 214.
Newton H. W.	1943	*Mon. Not. R. Astr. Soc.* **103,** 244;
	1944	*Ibid.* **104,** 4.
Newton H. W. and Jackson W.	1951	*Septième Rapport de la Commission pour l'Etude des Relations entre les Phénomènes Solaire et Terrestres* p. 107.
Van Sabben D	1952	*J. Atmosph. Terr. Phys.* **3,** 270.
Simon P.	1956	*Ann. Géophys.* **12,** 167.
Watson R. A.	1957	*J. Atmosph. Terr. Phys.* **11,** 59.

VOL. 14 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 296 to 298. Pergamon Press Ltd. Printed in Northern Ireland

On the origin of terrestrial particles from solar flares

W. H. WARD
Dominion Physical Laboratory, Department of Scientific and Industrial Research, Lower Hutt, New Zealand

(*Received* 19 *September* 1958)

Abstract—A mechanism which might lead to ejection of charged particles from the sun is suggested. It arises from the local cancellation of the magnetic field of a pair of sunspots by a general field.

LET us consider a pair of sunspots, of opposite magnetic polarity. The magnetic circuit will be closed by a loop external to the sun's surface, and round that field charged particles will circulate as a torus.

Let us now assume that the two sunspots are in an area of the sun that has a general magnetic field with a component along the line joining the spots. We need not assume a dipole field for the sun—the general field can come from the summation of local fields.

Fig. 1. - - - - - Local field of sunspot, —— —— general field and — — — sum of local and general fields; ——— difference of local and general fields.

Fig. 1 shows the mutual effect of the two fields. If the general field adds to the sunspot field, the torus will be compressed towards the sun's surface; if the two fields are opposed, the torus will be elongated. If the two opposing fields are of the same order of magnitude, and one varies relative to the other, the elongated torus will pulsate in and out. TELLER (1954) has dealt with the energies and motions of particles in the corona: "In this region the energy density due to the magnetic field is apt to be considerably greater than the energy density present

as hydrodynamic energy and thermal energy per unit volume carried by the material. Similarly, the Maxwell tensions due to the magnetic field are greater than the material pressures which would be available to drive a purely hydrodynamic motion. Numerically, for a field of 1000 G the magnetic energy density and the Maxwell stress will be $H^2/8\pi = 4 \times 10^4$ erg-cm^{-3}. Assuming a high velocity of 10^8 cm-sec^{-1} and a density of 10^8 protons and electrons per cm^3, one obtains a kinetic energy of only 0·8 erg-cm^{-3}. In order to estimate the thermal energy content and the pressure we assume a temperature $T = 10^{6°}$K. Such high temperatures are by no means unusual in the corona. With the same particle density as above one finds for the energy density 0·04 erg-cm^{-3} and for the pressure 0·027 erg-cm^{-3} $= 2{\cdot}7 \times 10^{-8}$ atm. Under these conditions the motion of the ionized material is almost completely determined by the magnetic lines of force." As the magnetic field moves from the closed path of the torus to the open path of the normal, the tenuous material of the corona will be swept along with it, and will end by being aimed at open space. That is, an ejection mechanism has been established. If the two fields cancel each other, the magnetic field of the sunspot may become effectively normal to the surface, even if only at a single height in the corona. KIEPENHEUER (1953) has published photographs that seem to be consistent with this mechanism.

VOL. 14 1959

The charged particles of the torus may have a high velocity. FERMI (1949). As they continue to travel out from the sun's surface some may be released from the magnetic field, through de-ionization or some other cause. Most of the particles that are released will no doubt fall back into the sun, colliding and raising the temperature as they go, but some may happen to be directed outwards, and may not lose their original velocity on being released from the magnetic field.

The particles which are not released will travel in a helix round a magnetic field which links one sunspot with some magnetic discontinuity other than the second sunspot. This may be the earth, in which case the particles will be directed to it. Their total velocity will be the same as that of the particles that are now free, but most of it will be taken up in circulation round the magnetic field; the forward velocity will correspond with the pitch of the helix.

$$p/r = 2\pi v/V$$

where p and r are the pitch and radius of the helix, V and v are the total and forward velocities.

ELLISON (1957) gives for the flare of 23 February 1956 a velocity of $c/3$ for the velocity of "cosmic ray protons" and of 900 km/sec for "magnetic storm particles (protons and electrons)." If the cosmic ray particles are neutral in transit between the sun and earth, or have been released from guidance by a magnetic field, we may take their velocity as giving the total velocity of the guided particles, and:

$$p/r = (2\pi \,.\, 900)/10^5 = 2\pi/100$$

In support of this elementary picture DENISSE (1952) may be quoted as showing that sunspots which emit geomagnetic corpuscles also emit radio noise storms, and WILD and ZIRIN (1956) and others, that spots which emit radio noise storms are associated with active looped prominences. WILD and ZIRIN claim that the

corpuscular streams are emitted when the prominences first move outwards and that the stream velocity greatly exceeds that of visible matter; this is believed to have a velocity of about 1000 km/sec in the direction of the earth. WILD *et al.* (1954) have shown that particles can leave the surface at a considerable angle to the vertical. Also PIGGOTT and ALLEN (1957) claim that there are no clear examples of the initial acceleration of the corpuscles in the sun, but frequent cases of deceleration. This is what one would expect if the particles were initially accelerated round the torus, but were decelerated after a cataclysmic change in the field which sustained them.

If it is a sound assumption that the magnetic field of a sunspot can "latch on" to the earth's field, it is to be expected that some (perhaps vanishingly small) change in the earth's magnetic field will occur at precisely the same time as the arrival of a radio noise storm.

REFERENCES

DENISSE J. J.	1952	*Ann. Géophys.* **8,** 55.
ELLISON M. A.	1957	*Neuvieme rapport de la commission pour l'etude des relations entre les phenomenes solaires et terrestres* p. 17. ICSU, Paris.
FERMI E.	1949	*Phys. Rev.* **75,** 1169.
KEIPENHEUER K. O.	1953	*The Sun* (Edited by KUIPER G. P.) pp. 413–414. University of Chicago Press.
PIGGOTT W. R. and ALLEN C. W.	1957	*Neuvieme rapport de la commission pour l'etude des relations entre les phenomenes solaires et terrestres* p. 79. ICSU, Paris.
TELLER E.	1954	*Rep. Prog. Phy.* **17,** 163.
WILD J. P., MURRAY J. D. and ROWE W. C.	1954	*Aust. J. Phys.* **7,** 439.
WILD J. P. and ZIRIN H.	1956	*Aust. J. Phys.* **9,** 315.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 299 to 324. Pergamon Press Ltd. Printed in Northern Ireland

Some possible effects caused by strong gyro-waves in the ionosphere—I

V. A. BAILEY
The University of Sydney, Sydney, Australia

Abstract—The possible increases of collision frequencies ν and electron densities n caused by a powerful extraordinary circular gyro-wave in the nocturnal lower E-region and in the day-time D-region are investigated on the basis of available laboratory experiments on electrons in air and data concerning these regions. The resultant effects of such increases on this wave and on other waves are then examined. It is found that some of these effects are remarkable in magnitude or in kind. Thus, notable self-distortion of gyro-waves and extremely large wave-interaction can be caused by such an extraordinary wave when the reduced equivalent electric force Z/p which it produces in the nocturnal lower E-region has the value 2·5 V/cm per mm Hg measured at room temperature.

Other possible detectable effects which may be caused at night when Z/p has a value between 2·5 and 8 include changes in the airglow and in parts of some trails of meteors and fast particles of solar or cosmic origin, changes in radar echoes from such trails and transient changes in the local magnetic elements.

1. INTRODUCTION

IT has been shown previously (BAILEY, 1937b, 1938) that an extraordinary circularly polarized gyro-wave radiated vertically from an orthogonal pair of horizontal dipoles suspended at a height $\lambda/4$ above the ground could cause observable wave-interaction when its power is as low as 50 W. This prediction was, in effect, confirmed by subsequent experiments in Italy (CUTOLO *et al.*, 1946) and in Australia (BAILEY *et al.*, 1952).

It was also shown (BAILEY, 1938) that when radiated vertically from an array of 400 similar pairs of orthogonal dipoles but with 500 kW of radiated power, such a gyro-wave could produce in the lower E-region a glow discharge which is "about fifty times as bright as the night sky on a moonless night." This prediction has not yet been tested.

With aerial arrays and radiated powers in between the extreme examples mentioned above it is possible for the gyro-waves to cause observable phenomena of other kinds (BAILEY and GOLDSTEIN, 1958).

It is therefore proposed to develop here an approximate theory of the effect of a strong extraordinary circular gyro-wave on the rates of attachment and recombination of electrons at different levels of the ionosphere and of the consequent changes in the electron densities at these levels. The resultant effects of such changes combined with accompanying changes in the collision frequencies, on other waves (including probing or "wanted" waves) will then be estimated. Other interesting possible effects will also be considered.

In outline the discussion will proceed as follows.

From the experiments on electrons in air carried out in Sydney (BAILEY, 1925) and Oxford (TOWNSEND and TIZARD, 1913) Table 1 is drawn up which gives the values of V the mean electron energy and b/p the reduced frequency of electron attachment for values of the reduced electric force Z/p in the range 0·5–2·5 V/cm per mm Hg. Reasons are given for provisionally applying these results to the air in the E- and D-regions.

From the table it is found that the linear relation (2.3) between b/p and Z/p holds in the given range of values of Z/p. In the absence of other reliable data this

VOL. 14 1959

relation is provisionally extrapolated to $Z/p = 0$. But the mathematical theory given later is independent of both (2.3) and its extrapolation.

The values or orders of magnitude of the required parameters relating to the level of the nocturnal ionosphere at 92 km are then estimated from available literature; the reasons for neglecting certain other parameters are also given.

Next the equivalent constant field of a circularly polarized gyro-wave of amplitude-E is defined.

The equations for the densities of the electrons and ions when $Z/p \leqslant 2{\cdot}5$ are considered in Section (4) and from them are derived appropriate approximate expressions for the densities n^+, n of the positive ions and electrons as functions of the time t and of the initial densities, and also for the densities in the steady state.

Expressions are then deduced for the time t_1 during which the electron density n varies linearly with time and the time t_2 after which n approaches within 5 per cent of its ultimate value.

The temporal variation of n after a long gyro-wave pulse is applied is determined in Section (6). It is found that initially the electron density increases at the rate of $100(b_0 - b_z)$ per cent per second during the time $t_{1z} = 0{\cdot}1/(b_z + c)$, where b_z is the value of b corresponding to the force Z, and that after the time t_{2z}, given below by equation (6.10), it approaches within 5 per cent of the ultimate density. The latter exceeds the initial density by $100(b_0 - b_z)/(b_z + c)$ per cent.

Thus, during the time t_{1z} for every 100 electrons the rate of attachment drops from $100b_0$ to $100b_z$ without appreciable compensating changes in the rates of change of the electron density through the processes of detachment, recombination or generation; this result is what might be expected from the fact that the change in the rate of attachment is much larger than the change in the rates of the other processes.

The consequences of the large changes in the collision frequency and the electron density caused in the nocturnal E-region by the gyro-wave are examined in Section (7). These include the effects produced on the gyro-wave's own attenuation and on the transverse propagation through the affected region of an ordinary wave of frequency 590 kc/s, under conditions like those current at Armidale.

Other effects which may be produced in the nocturnal E-region when Z/p has a value between 2·5 and 8 are then listed and discussed briefly.

The possible effects on wave-propagation and wave-interaction in the D-region, caused by the same extraordinary gyro-wave, are considered in Section (8).

The main results of the investigation are then summarized in Section (9).

2. Notation and Some Data

We shall here use as much as possible the well-established concepts and notation of Townsend and his collaborators. These are also given in the book by Healey and Reed (1940).

p = gas pressure (mm Hg).
T = gas temperature (°K).
Z = electric force (V/cm).
W = electron drift velocity (cm/sec).
k = Townsend's factor (electron temperature = kT).

V = mean energy of an electron (eV).
η = mean energy lost by an electron at a collision = ΔV.
u = r.m.s. random velocity of an electron (cm/sec).
l = mean free path of an electron (cm).
ν = collision frequency per second of an electron.
α_T = TOWNSEND's coefficient of ionization by collisions of electrons.
β = the coefficient of attachment of electrons.
$a = \alpha_T W$, the ionization frequency per second of an electron.
$b = \beta W$, the attachment frequency per second per electron.
c = rate of detachment of electrons per second from negative ions.
α, α_i = coefficients of recombination of electrons and negative ions, respectively, with positive ions.
q = rate of generation of electrons per cm^3 per sec.
Q_a = attachment cross-section (cm^2).
h = probability of attachment of an electron at a collision.
n, n^-, n^+, n^0 = number densities of electrons, negative ions, positive ions and neutral particles respectively.

The relations between some of our symbols and those of MASSEY and his collaborators are for convenience set out in the following table.

BAILEY	b	c	α	α_i	n	n^-	n^+	n°	q
MASSEY *et al.*	ηn	$\kappa n + \rho$	α_e	α_i	n_e	n^-	n^+	n	q

The values of a/p, b/p and h are obtained from the experimentally observed values of α_T/p, β/p and W by means of the formulae:

$$\frac{a}{p} = \frac{\alpha_T}{p} W; \qquad \frac{b}{p} = \frac{\beta}{p} W \tag{2.1}$$

$$h = 7 \times 10^{-16} (\beta/p) W^2/(Z/p) \tag{2.2}$$

For electrons in dry air at 288°K the observed values of k, β/p and W corresponding to different values of the "reduced force" Z/p are set out in the Table 1.

Table 1

Z/p	k	$(\beta/p) \times 10^3$	$W \times 10^{-5}$ (cm/sec)	V (eV)	b/p	At 92 km level		
						Z (mV/cm)	b	t_b (sec)
0·5	8·3	2·90	9·0	0·31	2610	0·83	4·44	0·22
1·0	15·2	1·85	12·5	0·56	2310	1·7	3·93	0·25
1·5	22·0	1·05	15·0	0·81	1580	2·55	2·69	0·37
2·0	27·5	0·55	17·5	1·02	960	3·4	1·63	0·61
2·5	33·0	0·35	19·5	1·22	680	4·25	1·16	0·86
0	1	∞	0	0·037	3300	0	5·61	0·18

VOL. 14 1959

The values of k and β/p are those found by BAILEY (1925) and the values of W are those found by TOWNSEND and TIZARD (1913). These values of k and W were later substantially confirmed by CROMPTON, HUXLEY and SUTTON (1953) and NIELSON and BRADBURY (1937) respectively.

The corresponding derived values of V the electron mean energy (eV) and of b/p the reduced attachment frequency, are given in the fifth and sixth columns of the table. V is calculated from the formula

$$V = k/27$$

since the mean energy of a molecule at 288°K is $\frac{1}{27}$ eV.

With regard to the tabulated values of β/p the following confirmatory evidence may be mentioned.

In Figs. 121 and 119B of his book, LOEB (1939) gives curves which represent the experimental values of respectively h and the "mobility" $k_e = W/760(Z/p)$ for electrons in air obtained by BRADBURY (1937). The corresponding results of BAILEY and of TOWNSEND and TIZARD respectively are also given there by means of dashed curves.

Fig. 119B shows that BRADBURY's values of W are about 40 per cent higher than those of TOWNSEND and TIZARD. Since the latter are in good agreement with the values of NIELSEN and BRADBURY (1937) it follows from equation (2.2) that BRADBURY's values of h in Fig. 121 must be reduced by the factor 1·96. When this is done the resulting curve is much closer to the curve which represents BAILEY values of h, although still above it by about 50 per cent.

LOEB's criticism that BAILEY's work was vitiated by impurities in his samples of gas is nullified by the fact that such impurities would place his curve *above* BRADBURY's, which we have just seen is not the case.

We may therefore conclude that the values of β/p given in Table 1 are substantially correct for air at 288°K and pressures down to 5 mm Hg.

In both investigations it was found that h depends on Z/p alone, i.e. h is independent of the air pressure down to about 5 mm Hg. Hence β/p is a function of Z/p alone. In the absence of reliable contrary evidence we may provisionally assume that this conclusion is applicable also when the air pressure is of the order of 2×10^{-3} mm as in the E-region or 4×10^{-2} as in the D-region.

The more recent results of HEALEY and KIRKPATRICK (1941) and others, on attachment in oxygen, are roughly consistent with these results on attachment in air.

The surprisingly low value of the attachment cross-section Q_a in oxygen obtained by BIONDI (1951) need not be considered here, for in a recent private communication Dr. BIONDI states that as a result of further studies in his laboratory he now regards this low value as erroneous.

If the data in Table 1 are used to plot the points $(Z/p, b/p)$ these are found to lie, within the limits of experimental error, on the straight line

$$\frac{b}{p} = 3300 - 1100\frac{Z}{p} \qquad (2.3)$$

in the range

$$0{\cdot}5 \leqslant Z/p \leqslant 2{\cdot}5$$

On certain theoretical grounds it is possible that b/p increases as the value of Z/p decreases below 0·5. So, if b_0/p is the value of b/p when $Z/p = 0$, we would then have $b_0/p > 2750$. In the absence of other reliable data we shall therefore not err much in taking

$$b_0/p = 3300 \tag{2.4}$$

which is a result of adopting the hypothesis that at 288°K the linear relation (2.3) may be extrapolated to $Z/p = 0$. We may then add the values given in the last row of Table 1.

In accordance with the results of numerous experiments on radio wave interaction and the views of many geophysicists it will be assumed here that the gas at the 92 km and 75 km levels is substantially the same as the air near the ground.

Since the temperatures T', T'' of the ionospheric regions near these levels (which we shall study) are both about 200°K it is now necessary to consider how to derive the appropriate values of the parameters required from their known values at the pressure p and temperature $T = 288°$K.

It is shown in the Appendix 1 that from the experiments of HALL (1955) and certain simple considerations we may infer the following approximate laws or rules of derivation when the force Z and molecular density n^0 have fixed values, i.e. with $p'/T' = p/T$.

When the mean energy of an electron is much greater than the mean energy V_0 of a molecule of the surrounding gas, i.e. $V \gg V_0$, and the electric force Z has a fixed value then, at the pressure p' and temperature T' of the gas the quantities V', W', η', h', $T'\nu'/p'$, $T'\beta'/p'$, $T'b'/p'$, $T'\alpha_T'/p'$, $T'a'/p'$, $T'k'$ are the same functions of the quantity $T'Z/p'$ as the respective functions which relate the quantities: V, W, η, h, $T\nu/p$, $T\beta/p$, Tb/p, $T\alpha_T/p$, Ta/p, Tk to the quantity TZ/p.

On applying these rules to determine the values of Z and b at the 92 km level which correspond to the data in Table 1 on Z/p and b/p, we obtain the following results: From the ARDC Model Atmosphere of MINZNER and RIPLEY (1956) we have at 92 km:

$$p' \doteqdot 1{\cdot}2 \times 10^{-3}\ \text{mm}, \qquad T' \doteqdot 200°\text{K},$$

hence $p = 1{\cdot}7 \times 10^{-3}$ when $T = 288°$K

and so, for $Z/p \geqslant 0{\cdot}5$

$$Z = 1{\cdot}7 \times 10^{-3} Z/p, \qquad b = 1{\cdot}7 \times 10^{-3} b/p$$

The values of Z and b so calculated for air at the 92 km level are given in the seventh and eighth columns of Table 1. The corresponding values of $t_b = 1/b$ the time constant of attachment are given in the ninth column.

We thus see that when the value of Z lies between 0·85 and 4·25 mV/cm the electron density n at this level changes through attachment alone by less than 1 per cent in 2 msec.

From the rules we also deduce that the corresponding mean electron energies V at the 92 km level are the same as those given in the fifth column of the table.

The value of b'_0/p', corresponding to $Z/p' = 0$, at the temperature 200°K cannot be determined by the rules of derivation given above, for we now have $V = V_0$. Since the mean energy at a collision is now 0·7 of that of a thermal electron in the laboratory experiments we have at present no alternative to the use of extrapolation. One procedure is to plot the points $(V, b/p)$ and extrapolate to $V = 0{\cdot}7V_0$ by means of the resulting curve. Such an extrapolation, consistent with the extrapolation of equation (2.3), yields values of b_0/p not greatly different from the value 3300 found earlier. We shall therefore assume that b_0 is independent of V_0, i.e. of T, in the range 200–288°K. Since p/T is constant we may therefore set

$$Tb_0/p = C \quad \text{(a constant)}$$

and so when $T = 288°$ we have by equation (2.4)

$$3300 = \mathrm{C}/288$$

From these results and equation (2.3) we obtain

$$\frac{Tb}{p} = C - 1100\frac{TZ}{p} \tag{2.5}$$

where

$$C = 0.95 \times 10^6 \tag{2.6}$$

On applying our rules of derivation to equation (2.5) under the condition $Z/p \geqslant 0{\cdot}5$ we obtain:

$$\frac{T'b'}{p'} = C - 1100\frac{T'Z}{p'} \tag{2.7}$$

when

$$T'Z/p' \geqslant 144$$

So at the 92 km and 75 km levels we have

$$b = 4750p' - 1100Z \tag{2.8}$$

in which, for convenience, we have dropped the prime from b'.

We note for future reference that, on account of equation (2.8), $b_0 - b$ is independent of the pressure of the region concerned.

It may now be noted that in recent years Bates and Massey (1951) have expressed the view (in effect) that for $p \ll 1$ mm Hg b_0/p diminishes with p in such a way that at the 92 km level it has a value which is thirty times smaller than ours. Their preferred estimates of b_0 in the ionosphere are, in their own view, not reliable; moreover, these estimates were found by them to be quite irreconcilable with Chapman's estimates of the conductivity required in the dynamo theory of magnetic variations. On the other hand our assumption that b_0/p is independent of p leads to values of b_0 at 92 km which are less difficult to reconcile with Chapman's estimates.

Other of their arguments (given elsewhere) against the importance of the process of electron-attachment in the E-region are based on the existence of persistent meteor trails and of a persistent nocturnal ionization in the E-region. Concerning the first we need only note Millman's (1952) remarks, that the basic

ionization in these meteor trails is "formed in an atmosphere containing an appreciable percentage of meteor atoms" and probably occurs "in a volume filled with heat energy at least 10^4 times as great as the energy of ionization," to see that no conclusions can be drawn from persistent trails about attachment in regions outside the trails. Concerning the second argument we should note that it was based on an assumption, that soon after sunset $q = 0$, which is now regarded as erroneous by KAISER, NICOLET and DUBIN (KAISER, 1955).

Values of c, the nocturnal rate of detachment of electrons, are not known. The experiments of BAILEY (1925) on electrons in air at 288°K indicated that when $Z/p \leqslant 2{\cdot}5$ detachment was quite undetectable; we may expect detachment to be even less in magnitude at 200°K. Consequently we shall here assume that:

$$c \ll b \tag{2.9}$$

With regard to the coefficients of recombination we shall here assume limiting values near those indicated by BATES (1952) and MASSEY (1950), namely:

$$10^{-8} < \alpha < 10^{-7} \tag{2.10}$$

$$5 \times 10^{-9} < \alpha_i < 5 \times 10^{-8} \tag{2.11}$$

The Australian and Italian experiments on gyro-interaction (BAILEY *et al.*, 1952; CUTOLO *et al.*, 1946; CUTOLO, 1950) indicate definitely that the influence of ions in the ionospheric regions concerned is quite negligible. Hence we may set

$$n^+ \ll 6 \times 10^4 n$$

Also the experimental conditions of gyro-interaction and the related theory combine to show that at the 92 km level soon after midnight the electron density n is of the order of 100 electrons/cm^3. We thus have

$$n \simeq 100, \qquad n^+ \ll 6 \times 10^6 \tag{2.12}$$

It follows from equations (2.12) and (2.10) that $\alpha n^+ \ll 0{\cdot}6$. Hence, by Table 1:

$$\alpha n n^+ \ll bn \tag{2.13}$$

i.e. the rate of electron recombination at the 92 km level after midnight is negligible compared with the rate of electron attachment.

The nocturnal values of q, the rate of generation of electrons/cm^3 per sec, considered as likely to exist in the main E-region (under the continual action of numerous meteors) by KAISER, NICOLET and DUBIN (KAISER, 1955), range from 2×10^{-4} to 20. For the 92 km level we may therefore assume that

$$q \ll 20 \tag{2.14}$$

Hence, by equation (2.12) and Table 1

$$q \ll bn \tag{2.15}$$

The value of the mean collision frequency ν_0 of an electron at the 92 km level will be assumed to lie in the range (NICOLET, 1952):

$$3 \times 10^5 \leqslant \nu_0 \leqslant 7 \times 10^5 \tag{2.16}$$

Summarizing our conclusions, we may say that for the 92 km level the degrees of certitude of our knowledge are as follows: The nature of the gas and its pressure p and temperature T are known approximately; the orders of magnitude of b, α, α_i, n and ν_0 are assumed known, but only upper limits can be assigned to the values of c, n^+, n^- and q.

Because of the apparent indefiniteness of this foundation to the theory given below we should mention here that several of the main conclusions turn out to be independent of the actual values of c, n^+, n^- and q, and largely independent of the values of α and α_i.

Since we shall be considering mainly phenomena which last less than 2 sec we shall neglect mass motions of the air. Also any contributions to the ionization at the 92 km and 75 km levels by ambipolar diffusion from and to other levels, will be regarded as included in the quantity q. We shall here also neglect the effects due to any constant or slowly varying electric fields X which may be present. In a more rigorous theory the effects of the earth's magnetic field H and of the electric fields X, on the diffusion, drift velocities, electron temperatures and the coefficients of attachment and electron recombination, should be taken into account, but this task will not be undertaken here.

3. Equivalent Field Z of a Circularly Polarized Gyro-wave with Rotating Electric Vector E

We define the equivalent field of a wave as that constant field Z which, in the steady state, does the same mean work $\bar{w}$ per collision of an electron in the medium as is done by the wave field over a period $2\pi/\omega$.

In Appendix 2 it is shown how to derive expressions for $\bar{w}$ for the extraordinary and ordinary polarized modes, of angular frequency ω, from the formula (21) of an earlier publication (BAILEY, 1937a). By equating these expressions to the expression for the mean work per collision done by a constant field Z, namely:

$$\bar{w} = (e^2/m)Z^2/\nu^2 \tag{3.1}$$

the formulae (A2.8) and (A2.9) for the equivalent fields are obtained. For an exact gyro-wave we have

$$\omega = \Omega$$

and then for propagation in a direction inclined at an angle θ to the terrestrial magnetic field these formulae yield:

$$Z_e = E \cos^2 \tfrac{1}{2}\theta \cdot \left(1 + \frac{2 \tan^2 \tfrac{1}{2}\theta}{1 + \Omega^2/\nu^2} + \frac{\tan^4 \tfrac{1}{2}\theta}{1 + 4\Omega^2/\nu^2}\right)^{\frac{1}{2}} \tag{3.2}$$

and

$$Z_0 = E \sin^2 \tfrac{1}{2}\theta \cdot \left(1 + \frac{2 \cot^2 \tfrac{1}{2}\theta}{1 + \Omega^2/\nu^2} + \frac{\cot^4 \tfrac{1}{2}\theta}{1 + 4\Omega^2/\nu^2}\right)^{\frac{1}{2}} \tag{3.3}$$

for the equivalent fields of the extraordinary and ordinary modes respectively.

When the wave is propagated along the magnetic field we have $\theta = 0$ and so $Z_e = E$. This result has been given as a theorem in an earlier publication (BAILEY 1938, Appendix 1).

When $$|\theta| < 45° \quad \text{and} \quad \nu^2 \ll \Omega^2 \tag{3.4}$$

we have the approximation:

$$Z_e = E \cos^2 \tfrac{1}{2}\theta \tag{3.5}$$

At Armidale, where experiments on gyro-interaction are current, we have $\theta \doteqdot 30°$ and so $Z_e = 0{\cdot}933\, E$. Also $\Omega = 10^7$ and, when E is small, $\nu = \nu_0 = 5 \times 10^5$ at the 92 km level. Hence we then have $\nu^2/\Omega^2 = 2{\cdot}5 \times 10^{-3}$. But when $E = 1{\cdot}82$ mV/cm the approximation (3.5) yields $Z_e = 1{\cdot}7$ which, by Table 1, yields $k = 15{\cdot}2$ for electrons at this level. Then the collision frequency rises to $\sqrt{(15{\cdot}2)} \times 5 \times 10^5$ approximately and so $\nu^2\,\Omega^2$ rises to $3{\cdot}8 \times 10^{-2}$. But since $\tan^2 \frac{1}{2}\theta = 0{\cdot}072$ the exact formula (3.2) yields a value of Z_e only 0·26 per cent higher than the approximate one (3.5). Thus equation (3.5) can here be safely used for much higher values of E.

The equivalent field for the corresponding ordinary mode is found from equation (3.3) to be $Z_0 = 0{\cdot}073E$ when E is so small that $\nu \doteqdot \nu_0$. Also when ν rises to $\sqrt{(15{\cdot}2)}\nu_0$ under the influence of a wave like that just considered with $E = 1{\cdot}82$, then equation (3.3) yields $Z_0 = 0{\cdot}131\, E$. Hence the effects due to an ordinary gyro-wave are always negligible compared with those due to an extraordinary one of the same amplitude.

VOL. 14 1959

4. The Equations for the Densities when $Z/p \leqslant 2{\cdot}5$

Under the influence of a gyro-wave with equivalent constant force Z the fundamental equations are:

$$Dn = (a - b)n + cn^- - \alpha n n^+ + q \tag{4.1}$$

$$Dn^- = bn - cn^- - \alpha_i n^- n^+ \tag{4.2}$$

$$n^+ = n + n^- \tag{4.3}$$

where $$D \equiv d/dt$$

and a is the ionization frequency of an electron per second.

We shall here limit the discussion to waves for which $Z/p \leqslant 2{\cdot}5$ when a is negligible compared with b. Accordingly, we shall take $a = 0$ in the following mathematical theory. Stronger waves for which $a > 0$ will be studied in Part 2.

When necessary, the fact that the quantities b and α are functions of Z will be indicated by writing them as b_z and α_z respectively. Since the thermal energies of the ions are not appreciably changed by the wave field, on account of the large masses and high frequency involved, we shall take the quantities c and α_i as independent of Z.

These equations are not exact, since they neglect several factors, including the non-uniformity of the gas and the densities, but we shall for simplicity follow the usual practice of neglecting these factors when only limited regions of the ionosphere are being considered.

Also, on account of uncertainties about the magnitudes of some of the constants involved we shall delay making further approximations as long as possible. This precaution unavoidably lengthens the discussion.

From equation (4.1) with $a = 0$ and equation (4.3) we have:

$$Dn = -(b + c)n + cn^+ - \alpha nn^+ + q \tag{4.4}$$

and by addition of equations (4.1) and (4.2) and use of equation (4.3):

$$Dn^+ = -n^+(\sigma n + \alpha_i n^+) + q \tag{4.5}$$

where

$$\sigma = \alpha - \alpha_i \tag{4.6}$$

The initial conditions are $n = N$, $n^+ = N^+$, $t = 0$.

In general it is not easy to obtain solutions of the system (4.4) and (4.5) in finite terms. So we shall begin by obtaining solutions for n and n^+ as power series in the time t by setting in these equations:

$$\left.\begin{aligned} n &= A_0 + A_1 t + A_2 t^2 + \dots \\ n^+ &= B_0 + B_1 t + B_2 t^2 + \dots \end{aligned}\right\} \tag{4.7}$$

and equating to zero the coefficients of t^0, t^1, etc., in each of them. This yields the following relations:

$$\left.\begin{aligned} A_1 &= q - (b + c)A_0 + cB_0 - \alpha A_0 B_0 \\ 2A_2 &= -(b + c)A_1 + cB_1 - \alpha(A_0 B_1 + A_1 B_0) \end{aligned}\right\} \tag{4.8}$$

$$\left.\begin{aligned} B_1 &= q - B_0(\sigma A_0 + \alpha_i B_0) \\ 2B_2 &= -\sigma(B_0 A_1 + B_1 A_0) - 2\alpha_i B_0 B_1 \end{aligned}\right\} \tag{4.9}$$

Also from the initial conditions we obtain

$$A_0 = N, \qquad B_0 = N^+ \tag{4.10}$$

If a steady state exists, then $Dn = 0$, $Dn^+ = 0$ and so the equations (4.4) and (4.5) yield the following equations for the steady densities $\bar{A}_0$, $\bar{B}_0$:

$$\left.\begin{aligned} -(b + c)\bar{A}_0 + c\bar{B}_0 - \alpha\bar{A}_0\bar{B}_0 + q &= 0 \\ -\bar{B}_0(\sigma\bar{A}_0 + \alpha_i\bar{B}_0) + q &= 0 \end{aligned}\right\} \tag{4.11}$$

It will be noted from equations (4.8) and (4.9) that the two conditions:

$$A_1 = 0, \qquad A_2 = 0$$

taken together imply that the steady state exists, for they yield equation (4.11). Hence, if n is constant over a sufficiently long period of time (which, by equation (4.21) below, may be as low as $1/\sqrt{[q\alpha_i]}$) we may conclude that the steady state is approximated.

On eliminating $\bar{A}_0$ between the equations (4.11) we obtain the following cubic in $\bar{B}_0$:

$$\alpha_i\alpha\bar{B}_0^3 + (\alpha_i b + \alpha c)\bar{B}_0^2 - \alpha_i q\bar{B}_0 - (b + c)q = 0 \tag{4.12}$$

This equation has only one positive root, so only one steady state can exist. In order to find a simple approximation to this positive root we may proceed as follows:

When $\alpha = \alpha_i$ the positive root of equation (4.12) is B_{01} where

$$B_{01} = \sqrt{(q\alpha_i^{-1})}$$

When $\alpha \neq \alpha_i$ we may use Newtonian iteration to obtain higher approximations with B_{01} taken as the first approximation. Thus we obtain the next approximation for the ion density in the steady state:

$$\bar{B}_0 \doteq \sqrt{(q\alpha_i^{-1})} - \frac{(\alpha - \alpha_i)\{q + c\sqrt{(q\alpha_i^{-1})}\}}{2\alpha_i b + 2\alpha c + (3\alpha - \alpha_i)\sqrt{(q\alpha_i)}} \tag{4.13}$$

Then, from equation (4.11) we obtain for $\bar{A}_0$ the value

$$\bar{A}_0 = \frac{q + c\bar{B}_0}{b + c + \alpha\bar{B}_0} \tag{4.14}$$

In the general state the infinite series (4.7) for the densities allow us to study the *initial* effects of a strong wave on the electron density and thus to obtain the relevant equations (6.12) and (6.19) in Section (6). But these infinite series are difficult to use for determining *the ultimate* effects of this wave. Consequently it is now necessary to introduce the conditions (2.13) and (2.15) in order to determine the ultimate effects without undue labour.

Under these two conditions equation (4.4) approximates to the following equation:

$$Dn = -(b + c)n + cn^+ \tag{4.15}$$

When studying the ultimate effects we are concerned with situations in which Dn is negligible, i.e. with conditions in which the ratio n/n^+ approximates to $c/(b + c)$. On account of equations (2.9), (2.10) and (2.11) we then have

$$|\sigma|n \ll \alpha_i n^+, \tag{4.16}$$

where

$$\sigma = \alpha - \alpha_i$$

We may then adopt the additional condition (4.16) and so use the resulting approximation to the equation (4.5) namely:

$$Dn^+ = -\alpha_i(n^+)^2 + q \tag{4.17}$$

The solutions of the system (4.15) and (4.17) can be used, when necessary, as first approximations from which to derive higher approximations to the system (4.4) and (4.5). Their validity for representing also initial effects can be tested by means of the solutions (4.7).

With N^+ as the initial value of n^+ the solution of equation (4.17) is

$$n^+ = B \tanh (kt + K_1) \tag{4.18}$$

or

$$n^+ = B \coth (kt + K_2) \tag{4.19}$$

where

$$\left.\begin{aligned} B &= \sqrt{(q\alpha_i^{-1})} \\ k &= \sqrt{(q\alpha_i)} \end{aligned}\right\} \tag{4.20}$$

$$K_1 = \tanh^{-1}(N^+/B), \qquad K_2 = \coth^{-1}(N^+/B)$$

VOL. 14 1959

We use (4.18) or (4.19) according as n is increasing or decreasing at $t = 0$.

The relaxation time of n^+ is τ^+, where

$$\tau^+ = 1/\sqrt{(q\alpha_i)} \tag{4.21}$$

These formulae show that n^+ is independent of b, c and α, and so also independent of Z when $Z/p \leqslant 2{\cdot}5$.

On account of equations (2.11) and (2.14) we have, from equation (4.21) $\tau^+ \gg 1000$ sec. It follows from this that during any interval of time less than 50 sec the positive ion density n^+ may be taken as constant, i.e.

$$n^+ = N^+ \quad \text{when} \quad 0 < t < 50 \text{ sec}$$

Therefore in the equation (4.15) we may now set the term cn^+ equal to the constant cN^+. This yields the solution:

$$\left.\begin{aligned} n = \bar{N}\{1 - \exp(-t/\tau)\} + N \exp(-t/\tau) \\ \text{for } t < 50 \text{ sec} \end{aligned}\right\} \tag{4.22}$$

where N is the initial value of n,

$$\bar{N} = N^+c/(b + c) \tag{4.23}$$

and

$$\tau = 1/(b + c) \tag{4.24}$$

Thus, n is approximately independent of α and α_i and depends on Z only through the attachment coefficient b.

Also we see that as t increases n tends asymptotically to the value $\bar{N}$ under the relaxation time τ when $\tau \ll \tau^+$, i.e. when

$$\sqrt{(q\alpha_i)} \ll b + c \tag{4.25}$$

This situation exists at the 92 km level on account of equations (2.11) and (2.14).

From equation (4.22) we have

$$n = N + A_1 t + A_2 t^2 + \ldots \qquad (0 < t \ll \tau) \tag{4.26}$$

where

$$\left.\begin{aligned} A_1 &= (b + c)(\bar{N} - N) \\ A_2 &= -\tfrac{1}{2}(b + c)^2(\bar{N} - N) \end{aligned}\right\} \tag{4.27}$$

We thus see that n varies linearly with t during the time t_1 where:

$$|A_2 t_1^2| = 0{\cdot}05|A_1 t_1|$$

i.e.

$$t_1 = 0{\cdot}1\tau \tag{4.28}$$

Also, n approaches within 5 per cent of its asymptotic value $\bar{N}$ when $t > t_2$ where

$$t_2 = \tau \log\left\{\frac{20\,|\bar{N} - N|}{\bar{N}}\right\} \ll \tau^+ \tag{4.29}$$

VOL
14
195

Although the formulae in this section have been derived primarily for application to the 92 km level, they are probably also valid for the 75 km level, as will be shown in Section (8).

In the next two sections we shall add the subscripts 0 or z to those symbols, such as b, α, n, n^-, n^+, etc., the values of which depend on the value of the perturbing force Z; the subscript 0 will indicate the unperturbed or normal values. The functional relations of b and α to Z will be taken as arbitrary, unless otherwise stated.

5. Temporal Variation of the Normal Densities

It is known, from observations at Watheroo and elsewhere, that on a normal quiet night between 0000 hours and about 0400 hours the electron density in the E-region decreases very slowly, often by less than 20 per cent in 1 hr, i.e.

$$\left|\frac{dn_0}{dt}\right| < 5 \times 10^{-5} n_0$$

It then follows from equation (4.15) that

$$(b_0 + c)n_0 = cn_0^+ \tag{5.1}$$

Since $\tau^+ \gg 10^3$ it follows that n_0^+ and n_0 remain constant during any interval of time less than 50 sec, i.e.

$$n_0 = \bar{N}_0 \qquad n_0^+ = N_0^+ \tag{5.2}$$

where

$$N_0^+ c = (b_0 + c)\bar{N}_0 \tag{5.3}$$

This result could also have been anticipated from the formulae (4.23) to (4.24), since the relaxation of n_0 would be practically complete in the time between sunset and midnight, which is very long compared with τ_0.

It should be noted that a relation equivalent to (5.3) has been held by Massey and his collaborators to apply to daytime densities. Consequently the present theory may be applicable also to the daytime regions.

6. Temporal Variation of the Electron Density Caused by a Gyro-wave Pulse

We suppose that a wave-pulse of constant equivalent amplitude Z is applied at $t = 0$ and ceases at $t = T$, where

$$2 \times 10^{-3} \ll T < 25 \text{ sec}$$

Initially the collision frequency ν rises from the normal value ν_0 to a value ν_z which does not change appreciably after 3 msec. This occurs because, as shown by experiments on gyro-interaction, the relaxation time for ν is less than 1 msec in the lower E-region. We may therefore assume, for simplicity, that at $t = 0$ the densities have their normal values $\bar{N}_0$, N_0^+ given under equation (5.2).

In the interval

$$0 \leqslant t \leqslant T < 25 \text{ sec}$$

the equations (4.22) to (4.29) assume here the following forms:

$$n_z = \bar{N}_z(1 - \exp(-t/\tau_z) + \bar{N}_0 \exp(-t/\tau_z) \tag{6.1}$$

$$\bar{N}_z = N_0{}^+c/(b_z + c) \tag{6.2}$$

$$\tau_z = 1/(b_z + c) \tag{6.3}$$

$$\sqrt{(q\alpha_i)} \ll b_z + c \tag{6.4}$$

$$n_z = \bar{N}_0 + A_{1z}t + A_{2z}t^2 + \ldots \tag{6.5}$$

$$A_{1z} = (b_z + c)(\bar{N}_z - \bar{N}_0)$$

$$A_{2z} = -\tfrac{1}{2}(b_z + c)^2(\bar{N}_z - \bar{N}_0)$$

$$t_{1z} = 0{\cdot}1\tau_z \tag{6.6}$$

$$t_{2z} = \tau_z \log\left\{\frac{20\,|\bar{N}_z - \bar{N}_0|}{\bar{N}_z}\right\} \ll \tau^+$$

From equations (6.2) and (5.3) we have

$$\bar{N}_z - \bar{N}_0 = \bar{N}_0 \frac{b_0 - b_z}{b_z + c} \tag{6.7}$$

and so

$$\left.\begin{aligned} A_{1z} &= \bar{N}_0(b_0 - b_z) \\ A_{2z} &= -\tfrac{1}{2}\bar{N}_0(b_z + c)(b_0 - b_z) \end{aligned}\right\} \tag{6.8}$$

$$t_{2z} = \tau_z \log\left\{\frac{20\,|b_0 - b_z|}{b_0 + c}\right\} \ll \tau^+ \tag{6.9}$$

Hence equation (6.5) yields

$$n_z = \bar{N}_0(1 + \rho_1 t + \rho_2 t^2 + \ldots) \tag{6.10}$$

where

$$\left.\begin{aligned} \rho_1 &= b_0 - b_z \\ \rho_2 &= -\tfrac{1}{2}(b_z + c)(b_0 - b_z) \end{aligned}\right\} \tag{6.11}*$$

Equation (6.10) shows that during the time $0{\cdot}1\tau_z$ the electron density increases

* A more exact calculation by means of the equations (4.7) to (4.10) shows that:

$$\rho_1 = b_0 - b_z + (\alpha_0 - \alpha_z)N_0{}^+$$

but with our assumed data the last term may be neglected.

linearly with time at the rate of $100(b_0 - b_z)$ per cent per sec and so in this time the density increases by

$$10 \cdot \frac{b_0 - b_z}{b_z + c} \quad \text{per cent} \tag{6.12}$$

Also, in the time t_{2z} given by equation (6.9) the electron density approaches within 5 per cent of the ultimate density $\bar{N}_z$, given by equation (6.7), which exceeds the initial density $\bar{N}_0$ by

$$100 \cdot \frac{b_0 - b_z}{b_z + c} \quad \text{per cent} \tag{6.13}$$

For example, when $Z/p = 2{\cdot}5$, we find by means of the Table 1 that the density linearly increases by nearly 40 per cent and ultimately increases by nearly 400 per cent. These increases will occur at the 92 km level after 0·086 and 2·3 sec respectively. Before they commence to occur the collision frequency will increase within 3 msec by a factor of about 10.

When the pulse length T is less than 25 sec the electron density at $t = T$ attains the value N_T where:

$$N_T = \bar{N}_z\{1 - \exp(-T/\tau_z)\} + \bar{N}_0 \exp(-T/\tau_z), \qquad (T < 25) \tag{6.14}$$

We now consider the variation of ν and n after the pulse ends.

During the first 3 msec the collision frequency ν decreases nearly to the normal value ν_0. Then during the following 25 sec the density n will vary according to the formula (4.22) with $b = b_0$, $\tau = \tau_0$ and $n = N_T$ at $t = 0$. Moreover, in that formula we must now set $\bar{N} = \bar{N}_0$ since n will return closely to its normal unperturbed value $\bar{N}_0$ after 25 sec; this result can also be deduced from equation (4.23) with $b = b_0$ combined with equation (5.3) and the fact that n^+ does not change appreciably from its value N_0^+ at the beginning of the period of 50 sec which we are considering.

Thus, after the pulse ceases n decays according to the formula:

$$n = \bar{N}_0\{1 - \exp(-t/\tau_0)\} + N_T \exp(-t/\tau_0) \tag{6.15}$$

where

$$\tau_0 = 1/(b_0 + c) \tag{6.16}$$

The subsequent theory is simplified when the pulse length T is much greater than τ_z. Thus, when

$$T \gg \tau_z \tag{6.17}$$

we have from equation (6.14) $N_T = \bar{N}_z$ and so

$$n = \bar{N}_0\{1 - \exp(-t/\tau_0)\} + \bar{N}_z \exp(-t/\tau_0) \tag{6.18}$$

where $\bar{N}_z$ is defined by equation (6.7), or

$$\bar{N}_z = \bar{N}_0 \frac{b_0 + c}{b_z + c}$$

It will be noted that equation (6.18) is the same as equation (6.1) with the subscripts z replaced by 0, and vice versa. Consequently all the formulae (6.2) to

(6.13) transformed in the same way can now be taken over. We thus conclude that the electron density now decreases linearly with time at the rate of $100(b_0 - b_z)$ per cent per sec during the time $0{\cdot}1\tau_0$ and that in this time it decreases by

$$10 \cdot \frac{b_0 - b_z}{b_0 + c} \quad \text{per cent} \tag{6.19}$$

Also, under the condition

$$\sqrt{(q\alpha_i)} \ll b_0 + c \tag{6.20}$$

the density approaches within 5 per cent of the normal unperturbed density $\bar{N}_0$ in the total time t_{20} given by

$$t_{20} = \tau_0 \log \left\{ \frac{20|b_0 - b_z|}{b_z + c} \right\} \tag{6.21}$$

From all these results we see that any experiment which allows the proportional rise and fall in time of the electron density near the 92 km level to be measured will yield separately the values of the three quantities $b_0 - b_z$, $b_z + c$ and $b_0 + c$. Any two of these will check the third and, if c is neglected, yield the values of b_0 and b_z.

Such an experiment would also immediately test the range of our hypothesis (2.8), for then

$$b_0 - b_z = 1100Z$$

and this formula leads to the remarkable conclusion that the initial percentage change of electron density per second is *independent* of the pressure and temperature in the region concerned and is equal to 110,000 times the equivalent electric force Z of the gyro-wave expressed in V/cm.

7. Some Effects Caused in the E-Region by a Long Pulse of Strong Gyro-waves

We shall now apply the foregoing theory and data to the 92 km level of the nocturnal E-region when it is subjected to a pulse of gyro-waves of such intensity that at this level

$$Z/p = 2{\cdot}5 \tag{7.1}$$

The corresponding value of the equivalent field Z is, by Table 1:

$$Z = 4{\cdot}35 \text{ mV/cm} \tag{7.2}$$

During the first 3 msec the mean energy of an electron will rise until it attains the value 1·22 eV. This corresponds to radiation of wavelength 10,100 Å. On a Maxwellian distribution of energies about 50 per cent of the electrons will have energies exceeding those which correspond to the strong OH and O_2 bands in the natural air-glow and about 30 per cent will have energies greater than the quantum of red radiation.

At the same time the electron collision frequency ν near this level will increase by a factor f of nearly 10. Then the electron density n will increase by a factor g which attains the value 1·43 at the time $t = 100$ msec and its ultimate value of nearly 5 at the time $t = 2300$ msec.

These changes in ν and n can produce large changes in the propagation of radio waves through the local lower E-region. The following table shows the changes in the quantities p_0^2, α, β, γ, γ_L, γ_T which enter into APPLETON's (1927) expression for the complex refractive index $M = \mu - i\chi$ in the magneto-ionic theory of wave propagation.

$Z/p = 0$	p_0^2	α	β	γ	γ_L	γ_T
$Z/p = 2{\cdot}5$	$p_0^2 g$	α/g	$\beta f/g$	γ/g	γ_L/g	γ_T/g

For either mode M is given by the formula:

$$M = \left(1 + \frac{1}{X + iY}\right)^{\frac{1}{2}} \tag{7.3}$$

where $X + iY$ is a well-known complicated expression. Convenient general approximations to $X + iY$ are obtainable by inspection from the formulae for M given in a recent publication (BAILEY, 1958). These approximations were derived by means of the following rational approximation to a radical:

$$\sqrt{(1 + s^2)} \doteqdot \frac{4 + 3s^2}{4 + s^2} \qquad (|s| < 1) \tag{7.4}$$

except when $s^2 \doteqdot -1$.

If we apply this to equation (7.3) it yields the following approximation to M:

when
$$X^2 + Y^2 > 1 \tag{7.5}$$

$$M \doteqdot \frac{X + 0{\cdot}75 + iY}{X + 0{\cdot}25 + iY} \tag{7.6}$$

From this we obtain the following approximations under the conditions (7.5):

$$\mu \doteqdot 1 + \frac{0{\cdot}5X + 0{\cdot}125}{X^2 + Y^2 + 0{\cdot}5X + 0{\cdot}062} \tag{7.7}$$

$$\chi \doteqdot \frac{0{\cdot}5Y}{X^2 + Y^2 + 0{\cdot}5X + 0{\cdot}062} \tag{7.8}$$

except when $X + iY \doteqdot -1$.

For an ordinary wave in transverse propagation

$$X = \alpha, \qquad Y = \beta$$

where

$$\alpha = -\omega^2/p^2{}_0, \qquad \beta = \nu_0\omega/p_0^2, \qquad p_0^2 = 3{\cdot}19 \times 10^9 n_0 \tag{7.9}$$

and ω is the angular frequency of the wave.

So in the region perturbed by the gyro-wave we have

$$X = \alpha/g, \qquad Y = \beta f/g \tag{7.10}$$

where

$$f \doteqdot 10, \qquad g \doteqdot 1{\cdot}43 \text{ or } 5$$

VOL. 14 1959

We shall now estimate the values of μ and χ for the "wanted wave" currently used to observe the effects produced at Armidale by a gyro-wave of moderate power. This wave has the frequency 590 kc/s and an angle of incidence $i_0 = 76°$ and is therefore propagated in the lower E-region nearly transversely to the terrestrial magnetic field. Since our purpose is primarily to obtain the orders of magnitude of the effects produced by a powerful gyro-wave we shall here neglect the influence on the wanted wave of both the magnetic field and the non-uniformity of the electron density.

Also, in order to simplify the arithmetical work we shall assign the following values to the unperturbed collision frequency ν_0 and electron density n_0:

$$\nu_0 = 5{\cdot}92 \times 10^5, \qquad n_0 = 86 \tag{7.11}$$

From these data and the formulae (7.7) to (7.10) we obtain the values of X, Y, μ, χ shown in the Table 2 for different times t elapsed after the pulse arrives at the 92 km level.

Table 2

t (msec)	0	3	100	2300
X	-50	-50	-35	-10
Y	8	80	56	16
μ	0·990	0·997	0·996	0·986
$\chi \times 10^3$	1·58	4·51	6·45	22·8
$185(\chi - \chi_0)$	0	0·542	0·900	3·93
ρ_{15}	1	0·582	0·407	0·0196
ρ_{15}^{-1}	1	1·72	2·46	51

Since the coefficient of absorption κ is equal to $\chi\omega/c$ a wanted wave which has unit amplitude on entering the E-region will emerge, after transversing the length s of its trajectory, with the amplitude A where

$$A = \exp(-\chi\omega s/c)$$

If A_0 and χ_0 relate to the unperturbed wave therefore

$$\rho_s \equiv A/A_0 = \exp\{(\chi_0 - \chi)\omega s/c\}$$

Under the conditions obtaining in the current experiments at the University of New England, Armidale, the horizontal projection of the trajectory of the wanted wave is about 70 km (Smith, 1957). For a conservative estimate of the effects produced by the perturbing gyro-wave we may therefore adopt the value 15 km for the length s in our formula for ρ_s. We thus obtain the following expression for the proportional change in the observed amplitude caused by the pulse of gyro-waves:

$$\rho_{15} = \exp\{-185(\chi - \chi_0)\} \tag{7.12}$$

where $\chi_0 = 1{\cdot}58 \times 10^{-3}$.

The values of $185(\chi - \chi_0)$, ρ_{15} and ρ_{15}^{-1} are given in the last three rows of the Table 2.

The fourth row shows that the refractive index μ increases slightly from the value 0·990 in the first 3 msec and then decreases slowly to 0·986. At all times its value is greater than the value of sin i_0 which is 0·970. So in spite of the large increase of n at the 92 km level the wave continues to be reflected from a higher level.

The last row shows that the observed amplitude is reduced by the factors 1·72, 2·46 and 51 after 3, 100 and 2300 msec respectively. These changes could be easily observed.

We may similarly estimate the increases with time of the observed amplitude after the pulse terminates.

The reduction of amplitude by the factor 1·72 is due solely to the increase of the collision frequency. The subsequent changes by the additional factors 1·43 and 29·6 are due to the change in the rate of electron attachment caused by the gyro-wave. Only the latter changes depend on the truth of our hypothesis that b/p is a function of Z/p alone.

It is not possible at present to consider how our conclusions would be affected by the views of BATES and MASSEY, since they do not appear to have published any estimates of the variation of b with Z or V.

If it is suggested speculatively that at the low pressures in the E-region the values of b for any value of Z is reduced by the same factor 30 which reduces our value of b_0 to theirs and if (2.9), (2.13) and (2.15) still hold true, our formulae (6.12) and (6.13) would then be changed only by the replacement of the term c by $c/30$; if then we neglect this term for lack of knowledge of its value, we find that the same values 2·46 and 51, of ρ_{15}^{-1} are obtained as before but that they occur in times which are thirtyfold longer than those given in the last two columns of Table 2; thus, even on this speculation we would observe in 3 sec a decrease of amplitude by the factor 1·43 due to the change in the rate of attachment caused by the gyro-wave.

VOL. 14 1959

The effect on its own attenuation caused by the extraordinary gyro-wave can be calculated in the same way. But a simpler method is to use the relevant formulae (81) and (82) for κ given in a previous publication (BAILEY, 1937b). From these we have

$$\mu \doteqdot 1$$

$$\kappa = 0{\cdot}0535(1 - \tfrac{1}{2}\sin^2\theta)n/\nu = 0{\cdot}0468n/\nu$$

Initially the absorption coefficient is κ_0, where

$$\kappa_0 = 0{\cdot}0468n_0/\nu_0 = 6{\cdot}8 \times 10^{-6}$$

and after the time t

$$\kappa = \kappa_0 g/f$$

The values of κ obtained in this way for different times t are given in Table 3.

Table 3

t (msec)	0	3	100	2300
$\kappa \times 10^5$	0·68	0·068	0·096	0·34
A	0·507	0·934	0·908	0·712

The corresponding factors of attenuation A in 1 km, given by

$$A = \exp(-10^5\kappa)$$

are shown in the last row.

We thus see that within 3 msec the gyro-wave penetrates much more easily to a higher level. The enhanced penetration then diminishes until after 2300 msec it remains constant but still notably greater than the initial penetration.

Besides these there are several other interesting effects which may be caused by sufficiently powerful pulses of gyro-waves. They are all briefly summarized as follows:*

7.1. *Radio propagation*

Self distortion of gyro-waves on the fundamental frequency f_1 and possible development of the harmonics of f_1.

7.2. *Wave-interaction*

(a) Increased absorption by the affected region of other kinds of "wanted waves", including perhaps cosmic noise.

(b) Increase or decrease of amplitude of the waves used with ionspheric recorders at frequencies about the critical frequency of the affected ionospheric level.

(c) Increase in the received intensity of v.h.f. waves scattered from the affected region.

7.3. *Night airglow*

Changes in the airglow, commencing with the near infra-red emissions.

7.4. *Trails or tracks of meteors, solar corpuscles and cosmic rays*

(a) Changes in those parts of the trails which lie in the region affected. Such changes may be observed visually or through their effects on radar echoes from meteors and aurorae, i.e. by interaction with radar waves.

(b) Making visible parts of the tracks of invisible micrometeors, solar corpuscles or cosmic ray particles.

7.5. *Magnetic variations*

(a) Changes in the local magnetic elements, both in the affected region and on the ground.

(b) Changes in the local gyro-frequency with consequent changes in the resonance effects observed in gyro-interaction and in the propagation of other waves passing through the affected region.

The effects under headings 7.1 and 7.2(a) and (b) may all be observable when $Z/p = 2{\cdot}5$. Some of the others may require $Z/p \geqslant 4$ or 8.† But there can be no doubt that at least 7.3 is observable when $Z/p \geqslant 11$ (Bailey, 1938, 1939).

* Some of these will be discussed more fully elsewhere.
† These will be discussed more fully elsewhere.

With regard to the effects on tracks, mentioned under Section 7.4, support for these possibilities is provided by the experiments of BEVAN (1949), at the University of Sydney, in which the tracks of α- and β-particles through Neon at 200 mm pressure were revealed by the glows produced about them by strong micro-waves.

All these effects will be subject to variations by resonance as the frequency of the disturbing wave is varied by about 10 per cent around the local normal gyro-frequency.

8. EFFECTS CAUSED IN THE *D*-REGION

The foregoing mathematical theory may be applied also to the 75 km level of the *D*-region (in the day-time) since the conditions with which this theory is associated can also be applied to that level.*

When the same powerful gyro-wave as that considered in Section (7) acts on the 75 km level the corresponding value of Z/p is about 0·17. This value lies outside the range of Table 1 and so the use of extrapolation is unavoidable.

The appropriate value of the electron density is also somewhat uncertain.

For these reasons we shall here give only the results of calculations which have been made on the following assumptions:

$$\nu_0 = 10^7, \qquad n_0 = 500 \tag{8.1}$$

From equation (2.3) we deduce that

$$\frac{b_0 - b_z}{b_z + c} = \frac{Z/p}{3 - Z/p + c/1100} \tag{8.2}$$

and so from equation (6.13) we infer that the ultimate increase of electron density is by 6 per cent. From equations (2.8), (6.3) and (6.9) we find that this increase is closely approached within about 135 msec.

From the experiments on air of CROMPTON *et al.* (1953) we may infer that the collision frequency ν increases by a factor of about 6; this occurs within 0·2 msec after the pulse commences.

With a weak gyro-wave $\nu \doteqdot \nu_0$ and its amplitude decreases in 1 km by 20 per cent. With our strong gyro-wave $\nu \doteqdot 6\nu_0$ and then its amplitude decreases in 1 km by only 4 per cent. This is another example of the "non-linear" propagation found in Section (7). As a result the powerful gyro-wave penetrates much more deeply into the ionosphere than would a weak one.

The effects produced on the propagation, through the region near the 75 km level, of an ordinary wanted wave of not too low a frequency are as follows: when the refractive index μ is initially sufficiently close to 1 it does not change much. On the other hand, if the normal attenuation factor in the affected region is of the order of exp (−1) then within 0·2 msec this changes to exp (−6) and after 135 msec it changes by an additional factor of exp (−0·36), i.e. by the factor 0·7; both these reductions in amplitude should be observable by means of appropriate receiving equipment. Such effects may perhaps be observed on cosmic noise.

* The conditions (2.9), (2.13), (2.15) hold true (if we neglect the effects of the ions by comparison with those of the electrons) and the conditions (2.10), (2.11) may be assumed true at 75 km.

9. Summary and Conclusion

The main results of our investigation may be summarized as follows.

A theory is developed of the effects produced near the 92 km level of the nocturnal E-region and the 75 km level of the D-region by a pulse of extraordinary gyro-waves of such field intensity that the values of the equivalent reduced force Z/p at these levels are 2·5 and 0·17 V/cm respectively.

The theory is based as closely as possible on the most reliable experimental data about the behaviour of free electrons in air and on the present state of knowledge about the parameters which relate to the two levels. Where possible the effect of the existing divergence of opinion on the rates of attachment at pressures of the order of 0·002 mm Hd is examined.

It is found that near the 92 km level the collision frequency ν increases tenfold and the electron density n increases fivefold in less than 2·3 sec, while near the 75 km level ν increases sixfold and n increases by 6 per cent in less than 0·135 sec.

The consequent changes in the propagation characteristics of a transverse ordinary wave of frequency 590 kc/s which traverses 15 km of the region near the 92 km level are given in the Table 2. The last row of this table shows that the observed wave-amplitude decreases by the factors 1·72, 2·46 and 51 in the times 3, 100 and 2300 msec, respectively, after the pulse enters the region.

The effects on its own attenuation in 1 km caused by the pulse near the 92 km level, are given in the Table 3. This table shows that by reason of its great intensity the pulse can, within 3 msec, increase considerably its penetration into the E-region.

Under the influence of the pulse the mean energy of an electron at the 92 km level rises to 1·22 eV. This is equivalent to a quantum of energy of radiation of wavelength 10,100 Å, i.e. radiation close to the OH and O_2 bands of the natural airglow.

By the use of pulses with this field intensity, or with up to four times greater intensities, it may be possible to produce observable changes in the local airglow, in the trails of meteors and tracks of different fast particles which enter the lower E-region, and in the local magnetic elements. The probability of producing such observable changes increases rapidly with further increase of the intensity.

All these effects are subject to resonance variations as the pulse frequency varies about the local gyro-frequency.

The theory also shows that it is possible to use such a pulse to determine by experiment the rates of attachment of electrons near the 92 km and 75 km levels of the ionosphere as well as other data about the collisions of electrons with the neutral particles near these levels. In this way it is possible to resolve certain important questions which are at present the subject of speculation. These include the rates of attachment and detachment, the ratio of negative ions to electrons, the relative contributions of the E- and D-regions to the production of magnetic variations by dynamo action, the contribution of the E-region to the airglow, the existence of meteors and corpuscles of solar origin which are at present unperceived and perhaps also the causes of abnormalities such as sporadic-E ionization.

On account of the large increases caused in ν alone it will also be possible to simplify studies of wave-interaction by using vertically incident wanted waves

VOL
14
195

and so having all the necessary facilities at the one place instead of at three widely separated stations.

All these results taken together emphasize the importance of carrying out experiments with powerful gyro-waves of the kind considered here. The large initial cost would be amply repaid by having in existence a potent instrument with relatively low running expenses which would throw light on a large variety of processes in the ionosphere and perhaps, in due course, lead to improvements in radio-communication.

Thus, the erection in some centre of a suitable large aerial array of, say, 80 or 160 horizontal dipoles connected to a generator of 500 or 1000 kW power may be almost as worthy of support as the erection of a large radio-telescope.

Acknowledgements—I should like to acknowledge the valuable help in the preparation of this paper which has been given by Drs. R. A. SMITH and K. LANDECKER of the University of New England. Thanks are also due to Professor J. M. SOMERVILLE for making possible this and other assistance.

APPENDIX 1

First we assume that, if Q is the effective momentum cross-section of a molecule towards an electron of mean energy V, then

$$Q \text{ is independent of temperature } T \tag{1}$$

Since $\nu = u/l$ and $Qln^0 = 1$, therefore

$$\nu/n^0 = uQ \tag{2}$$

Since $W = Ze/m\nu$, therefore

$$W = (e/m)(Z/n^0)/uQ \tag{3}$$

In the steady state we must have

$$ZeW = \nu\eta$$

Hence by (3) and (2)

$$\eta = mW^2 = (e^2/m)(Z/n^0)^2/u^2Q^2 \tag{4}$$

Since

$$\tfrac{1}{2}mu^2 = Ve \tag{5}$$

therefore

$$mW^2 = \eta = \tfrac{1}{2}e(Z/n^0)^2/VQ^2 \tag{6}$$

Also, since

$$p = Kn^0T, \tag{7}$$

where K is Boltzmann's constant, we have

$$Z/n^0 = KTZ/p \tag{8}$$

Since the energy η lost at a collision is a function of V and T alone it follows that Z/n^0 is a function of V and T alone.

Now the experiments of HALL (1955) on electrons in N_2 and H_2 show that in the range of temperatures 83°K to 288°K and for fixed values of Z and n^0 the energy V is independent of the temperature T when $V \gg V_0$, where V_0 is the mean energy of a molecule of the gas, i.e.

$$V_0e = \tfrac{3}{2}KT \tag{9}$$

If we now assume that the same behaviour is true of electrons in air we may conclude that when

$$V \gg V_0 \tag{10}$$

Z/n^0 is a function of V alone, i.e. V is a function of Z/n^0 alone.

It follows from equations (1), (6), (5) and (2) that also η, W, u and ν/n^0 are functions of Z/n^0 alone.

Since the experiments on attachment in air show that at a fixed temperature the probability of attachment h depends on Z/p alone, it follows that h depends on V and V_0 alone. Therefore when equation (10) holds true h depends on V alone, i.e. on Z/n^0 alone.

Since νh is equal to the attachment frequency b and $b = \beta W$ it follows that b/n^0 and β/n^0 are functions of Z/n^0 alone.

Evidently the probability of ionization i of a molecule by an electron under the condition (10) depends only on V, i.e. on Z/n^0 alone. Since $\nu i = a$ and $a = \alpha_T W$ it follows that a/n^0 and α_T/n^0 are functions of Z/n^0 alone.

Lastly, from the definition of TOWNSEND's factor k as V/V_0 and from equations (9) and (10) we see that Tk is a function of Z/n^0 alone.

On using equation (8) we may therefore conclude that

$$\left.\begin{array}{c} V,\ W,\ \eta,\ h,\ T\nu/p,\ T\beta/p,\ Tb/p,\ T\alpha_T/p,\ Ta/p,\ Tk \\ \text{are all functions of } TZ/p \text{ alone} \end{array}\right\} \tag{11}$$

APPENDIX 2

If $\mathbf{h}$ is a unit vector in the direction of the local terrestrial magnetic field $\mathbf{H}$ and $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$ are three mutually perpendicular unit vectors with $\mathbf{n}_3$ in the direction of propagation of a circularly polarized wave of angular frequency ω and amplitude $\mathbf{E}$, we have

$$\mathbf{H} = H\mathbf{h}$$

and

$$\mathbf{E} = E_1 \cos \omega t \,.\, \mathbf{n}_1 + E_2 \sin \omega t \,.\, \mathbf{n}_2$$

where

$$E_1 = |E_2| = E$$

We also introduce the quantities Z_1, Z_2 defined by

$$Z_1^2 = \tfrac{1}{2}E_1^2, \qquad Z_2^2 = \tfrac{1}{2}E_2^2 \tag{1}$$

If θ is the angle between $\mathbf{h}$ and $\mathbf{n}_3$, $\quad c_{hn} = \mathbf{hn}$, $\quad \mathbf{s}_{hn} = \mathbf{h} \times \mathbf{n}$, $\quad g_{12} = \mathbf{h} \,.\, \mathbf{n}_1 \times \mathbf{n}_2$

and, for convenience, we take $\mathbf{n}_1$ in the plane containing $\mathbf{h}$ and $\mathbf{n}_3$, we have

$$\left.\begin{array}{ll} c_{hn_1}^2 = \sin^2\theta, & s_{hn_1}^2 = \cos^2\theta \\ c_{hn_2}^2 = 0, & s_{hn_2}^2 = 1, \qquad g_{12} = \cos\theta \end{array}\right\} \tag{2}$$

The angular gyro-frequency Ω is now given by

$$\Omega = -He/mc$$

From the formula (21) given by BAILEY (1937a) for the mean work $\overline{w}$ done by the wave per collision, in the steady state (for which $\overline{w}_0 = 0$), we have

$$\overline{w}/(e^2/m) = \sum_n Z_n^2(A_n \tfrac{1}{2}s_{hn}^2 + B_n c_{hn}^2) + Z_1 Z_2 C_n \mathbf{h} \,.\, \mathbf{n}_1 \times \mathbf{n}_2 \tag{3}$$

where

$$A_n, C_n = \frac{1}{\nu^2 + (\omega_n - \Omega)^2} \pm \frac{1}{\nu^2 + (\omega_n + \Omega)^2}$$

$$B_n = \frac{1}{\nu^2 + \omega_n^2}$$

Here the two plane components of the circular wave have the same frequency ω, so $\omega_1 = \omega_2 = \omega$. Also using equations (1) and (2) we thus obtain

$$\begin{aligned} \overline{w}/(e^2/m) &= \tfrac{1}{2}E^2(\tfrac{1}{2}A(\cos^2\theta + 1) + B\sin^2\theta) \pm \tfrac{1}{2}E^2 C\cos\theta \\ &= \tfrac{1}{4}E^2(d(1 \pm \cos\theta)^2 + s(1 \mp \cos\theta)^2) + \tfrac{1}{2}E^2 B\sin^2\theta \end{aligned}$$

where

$$d = \frac{1}{\nu^2 + (\omega - \Omega)^2}, \qquad s = \frac{1}{\nu^2 + (\omega + \Omega)^2}, \qquad B = \frac{1}{\nu^2 + \Omega^2} \tag{4}$$

On selecting the signs appropriate to the two modes we obtain the following formulae:
For the extraordinary mode

$$\overline{w} = (e^2/m)E^2(d \,.\, \cos^4 \tfrac{1}{2}\theta + s \,.\, \sin^4 \tfrac{1}{2}\theta + \tfrac{1}{2}B \,.\, \sin^2 \theta) \tag{5}$$

and for the ordinary mode

$$\overline{w} = (e^2/m)E^2(d \,.\, \sin^4 \tfrac{1}{2}\theta + s \,.\, \cos^4 \tfrac{1}{2}\theta + \tfrac{1}{2}B \,.\, \sin^2 \theta) \tag{6}$$

For a constant field Z alone we have the two relations:

$$\nu\bar{w} = ZeW$$

and

$$W = Ze/m\nu$$

and therefore the relation

$$\bar{w} = (e^2/m)Z^2/\nu^2 \tag{7}$$

Hence, by definition the equivalent fields Z_e, Z_0 of the extraordinary and ordinary modes are given respectively by

$$Z_e = E\nu(d \,.\, \cos^4 \tfrac{1}{2}\theta + s \,.\, \sin^4 \tfrac{1}{2}\theta + \tfrac{1}{2}B \sin^2 \theta)^{\frac{1}{2}} \tag{8}$$

and

$$Z_0 = E\nu(d \,.\, \sin^4 \tfrac{1}{2}\theta + s \,.\, \cos^4 \tfrac{1}{2}\theta + \tfrac{1}{2}B \sin^2 \theta)^{\frac{1}{2}} \tag{9}$$

Note Added in Proof

The new experimental results on the attachment of electrons in oxygen at pressures exceeding 7 mm Hg reported by L. M. Chanin, A. V. Phelps and M. A. Biondi, at the Gaseous Electronics Conference on October 23, 1958, cannot be extrapolated to the air at the 92 km level, for this would require the errors in the reported measurements of β/p to be less than 1 part in 10^4. In a private communication Dr. Biondi agrees with this conclusion.

References

Appleton E. V.	1927	*URSI Reports.* Washington.
Bailey V. A.	1925	*Phil. Mag.* **50,** 825.
Bailey V. A.	1937a	*Phil. Mag.* **23,** 774 (formula 21)
Bailey V. A.	1937b	*Phil. Mag.* **23,** 929.
Bailey V. A.	1938	*Phil. Mag.* **26,** 425.
Bailey V. A.	1939	*A.W.A. Tech. Rev.* **4,** 1.
Bailey V. A.	1958	*J. Atmosph. Terr. Phys.* **12,** 118.
Bailey V. A., Smith R. A., Landecker K., Higgs A. J. and Hibberd F. H.	1952	*Nature, Lond.* **169,** 911.
Bailey V. A. and Goldstein L.	1958	*J. Atmosph. Terr. Phys.* **12,** 216.
Bates D. R.	1952	*Proc. Conf. Ionosph. Phys. A, July* 1950 p. 225. A. F. Camb. Res. Center.
Bates D. R. and Massey H. S. W.	1951	*J. Atmosph. Terr. Phys.* **2,** 1.
Bevan A. R.	1949	*Nature, Lond.* **164,** 454.
Biondi M. A.	1951	*Phys. Rev.* **84,** 1072.
Crompton R. W., Huxley L. G. H. and Sutton D. J.	1953	*Proc. Roy. Soc.* A **218,** 507.
Cutolo M., Carlevaro M. and Gherghi M.	1946	*Alta Frequenza* **15,** 111.

CUTOLO M.	1950	*Nature, Lond.* **166,** 98.
HALL B. I. H.	1955	*Aust. J. Phys.* **8,** 551.
HEALEY R. H. and REED J. W.	1940	*The Behaviour of Slow Electrons in Gases* A. W. A. Sydney.
HEALEY R. H. and KIRKPATRICK C. B.	1941	*A.W.A. Tech. Rev.* **5,** 155.
KAISER T. R.	1955	*Meteors* (KAISER p. 119, NICOLET M. p. 99, DUBIN M. p. 111). Pergamon Press, London.
LOEB L. B.	1939	*Fundamental Processes of Electrical Discharge in Gases* Chapter VI. Wiley, New York.
MASSEY H. S. W.	1950	*Negative Ions* p. 97. Cambridge University Press.
MILLMAN P. M.	1952	*Proc. Conf. Ionosph. Phys. A, July* 1950 p. 204. A.F.C. Res. Center.
MINZNER R. A. and RIPLEY W. S.	1956	*ARDC Model Atmosphere.* A.F.C. Res. Center.
NIELSEN R. A. and BRADBURY N.	1937	*Phys. Rev.* **51,** 69.
NICOLET M.	1952	*Sci. Report. No.* 32. Ionosph. Res. Lab., State College, Penna.
SMITH R. A.	1957	Ph. D. Thesis, University of Sydney.
TOWNSEND J. S. and TIZARD H. T.	1913	*Proc. Roy. Soc.* **88,** 336.

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 325 to 337. Pergamon Press Ltd. Printed in Northern Ireland

Absolute photometry of the aurora—I

The ionized molecular nitrogen emission and the oxygen green line in the dark atmosphere

M. H. REES*

Boulder Laboratories, National Bureau of Standards and University of Colorado, Boulder, Colorado

(*Received* 30 *August* 1958)

Abstract—A low latitude aurora was measured photometrically. Three prominent emission features, λ3914 band of 1.N.G. N_2^+, λ5577 of O*I* and λ6300–6364 of O*I*, were isolated with narrow pass band filters and recorded continuously throughout the night in the meridian sweep. Absolute calibration was carried out, and corrections for the scattered light and atmospheric extinction were made.

The absolute emission intensity for λ3914 and λ5577 is given at a selected zenith distance. The ratio of instantaneous intensities I(O)5577 to I(O)3914 was found to be essentially constant over an intensity range of three powers of ten and to be independent of the type of auroral form.

The intensities and correlations obtained are consistent with the following excitation mechanisms. In the dark atmosphere the N_2^+ is excited from ionizing collisions of incoming particles with atmospheric N_2. It is concluded that the O*I* λ5577 line results from secondary electrons, arising from the primary ionization process above, colliding inelastically with oxygen atoms.

VOL. 14 1959

1. INTRODUCTION

THE desirability of carrying out auroral observations from which the absolute intensity of the various emission features can be determined has been realized for some time. On the basis of this experimental data, from which relative intensities are easily determined, considerable progress may be made in establishing the excitation mechanism primarily responsible for the auroral emissions. In this series of papers we report on results obtained from a photometric study of the aurora of 1–2 March 1957 as observed from the Fritz Peak Observatory in Colorado (latitude 39°55′N longitude 105°29′W, elevation 2751 m).

The photometer incorporated three telescopes equipped with narrow band interference type filters and auxiliary glass filters. Peak transmission and band width at half transmission of the filter systems, as determined by a high resolution laboratory spectrophotometer, were the following: λ3914, 27 Å wide; λ5577, 17 Å wide; and λ6300, 100 Å wide. These wavelengths correspond to three of the most prominent auroral emissions, the "forbidden" atomic oxygen lines at λ6300/6364 and λ5577, and the $v' = 0$ to $v'' = 0$ vibrational band of the First Negative system of ionized molecular nitrogen at λ3914. In view of the high intensity of these emission features and the narrow pass band of the filter systems employed it is permissible to neglect other radiations present in the respective wavelength intervals of the filters: (a) other auroral emission bands, (b) light of astronomical origin, i.e., integrated starlight and zodiacal light, and (c) the scattered light arising from the two preceding emission sources. The usual airglow emissions of the oxygen atom are, of course, accepted by the respective filters; however, the intensity of the normal airglow is quite low in comparison to auroral intensities.

* Now at the Geophysical Institute, University of Alaska

The photometer described vertical circle sweeps in the meridian alternately from north to south and from south to north. The duration of each sweep was 3 min. A shutter closed at the end of each sweep for one minute, yielding a zero reading with each sweep; the basic time interval of the observations was therefore four minutes. This instrument was run continuously from 2000 MST (Mountain Standard Time) to 0530 MST. The data were displayed on continuously running graphic recorders. The photometer, originally designed for airglow research by Marlow and Pemberton (1949), is described in some detail by Roach and Pettit (1951). Sample sweeps of the auroral radiations recorded are shown in Fig. 1. These sweeps illustrate the general trend of the display as observed from Fritz Peak, namely, that the "green" and "blue" emissions were concentrated primarily toward the northern horizon, while the "red" emission extended well over the entire northern half of the hemisphere, and frequently spilled over into the southern half.

2. Analysis of the Observations

The steps followed in obtaining the absolute emission intensity from the raw galvanometer deflexions were the following: (a) absolute calibration of the telescopes and (b) correction of the data for atmospheric scattering and extinction.

(a) Absolute calibration has been carried out by utilizing the deflexions due to a bright star. At the beginning of the night, when the aurora was still relatively faint, and the sensitivity setting of the system correspondingly high, the star α Canis Majoris crossed the meridian and gave a well defined deflexion on all three records. By a procedure described in detail by Rees (1958) a calibration factor yielding the surface brightness per unit galvanometer deflexion was obtained. All the data are presented in the absolute intensity unit, the kilorayleigh (kR) equivalent to $4\pi B$, where B is the angular surface brightness in units of 10^9 quanta-cm^{-2}-sec^{-1} sterad^{-1}. This photometric unit for the aurora has recently been suggested by Hunten, Roach and Chamberlain (1956).

(b) Correction for atmospheric scattering and extinction due to the terrestrial atmosphere situated between the region of auroral emission and the observer may be made by applying the equation of radiative transfer. For a point source at infinity (the sun) illuminating a plane parallel atmosphere a solution in closed form has been obtained by Chandrasekhar (1950). Ashburn (1953) integrated an infinite number of point sources over a hemisphere thereby obtaining the solution for a uniformly illuminated atmosphere. Such a solution is applicable to the transfer problem of the airglow which is more or less a uniform glow over the entire sky. The aurora is characterized by a non-uniform emission over only a portion of the sky as well as by a rapid change with time. An exact solution of the transfer problem involving an aurora is clearly impracticable and judicious approximations must therefore be used.

The formula used for making this correction is

$$I_\nu = I_\nu(\mathrm{O}) \exp(-\tau_\nu) + J_\nu\{1 - \exp(-\tau'_\nu)\}, \qquad (1)$$

where I_ν is the observed intensity at frequency ν, $I_\nu(\mathrm{O})$ is the corresponding intensity above the intervening atmosphere, τ_ν is the extinction coefficient due

VOL.
14
1959

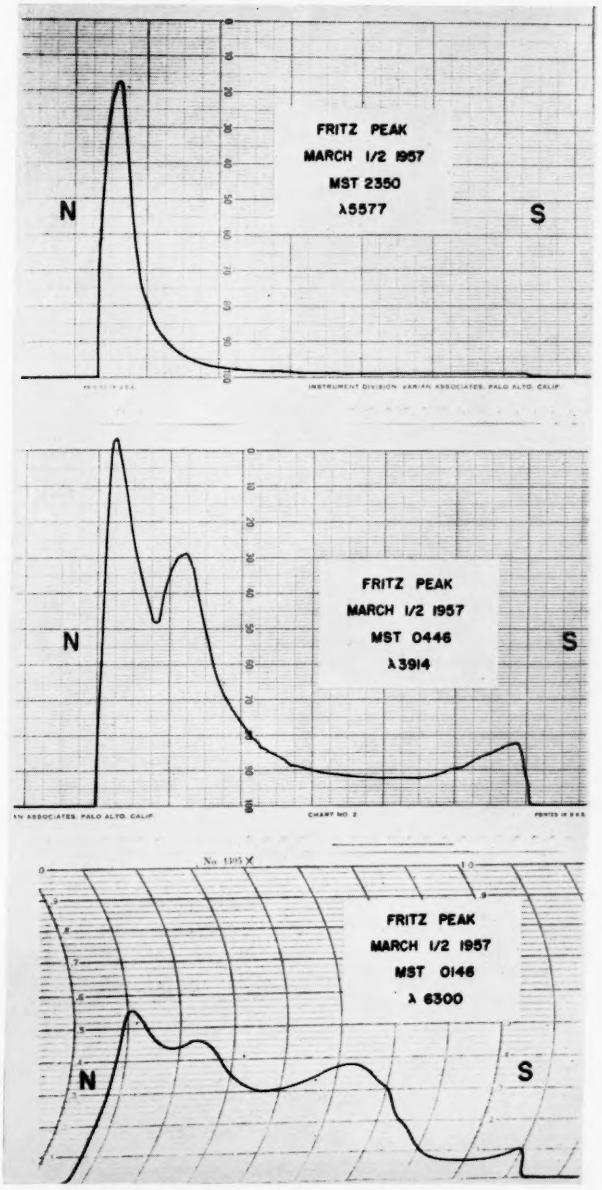


Fig. 1. Typical meridional sweeps of the auroral radiation recorded photometrically through selective filters.

VOL
14
195

to scattering and absorption, τ_ν' is the extinction coefficient for scattering only, and J_ν is an average intensity from which the scattered light is estimated.

Equation (1) was established by BARBIER (1944) primarily for application to airglow analysis. The important assumption implicit in equation (1) is that J_ν is independent of position in the atmosphere.

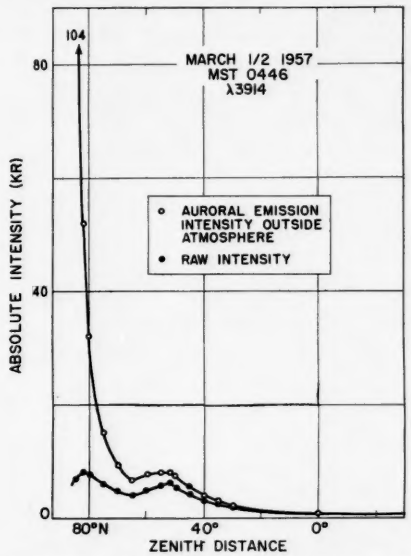


Fig. 2. Auroral intensity of the λ3914 emission before and after corrections for atmospheric extinction and scattering have been applied.

The angular dependence of the scattered light was taken into account using the Rayleigh phase function for molecular scattering:

$$1 + \cos^2 \Theta \tag{2}$$

where Θ is the angle between the incident and the scattered radiation. The areal extent and position of the aurora were estimated from visual observations. Details of the analysis are given by REES (1958).

In a photometric study of the airglow and the aurora at College, Alaska, BARBIER and PETTIT (1952) used a similar approach to correct for the scattered light and extinction. In their situation the aurora extended over a considerable portion of the sky; they therefore assumed J_ν to be constant all over the sky and obtained its value by a suitable treatment of the observational data. In this study we also determine J_ν from the observations; however, we use a different approximation: we assume that there is no auroral emission present near the southern horizon at $Z = 85°$S. Thus the intensity measured at $Z = 85°$S through the λ3914 filter is attributed entirely to scattered light from the aurora in the north. Although an emission at λ3914 has been reported as an airglow feature probably due to bombardment by cosmic ray protons (CHAMBERLAIN and OLIVER, 1953; MEINEL, 1953) its intensity is sufficiently low in comparison to the auroral scattered light component near the southern horizon to allow us to neglect any

VOL. 14 1959

airglow contribution in the analysis. In the case of the oxygen lines, in addition to scattered auroral light we also account for airglow and scattered light due to airglow, using average airglow intensities. Internal consistency in the results obtained for all measured emissions at all pertinent zenith distances lend support to the validity of the assumptions made.

In Figs. 2, 3 and 4 a comparison is made between the raw intensity after calibration only and the auroral emission intensity above the atmosphere. The striking effect of increasing atmospheric extinction toward the horizon and its

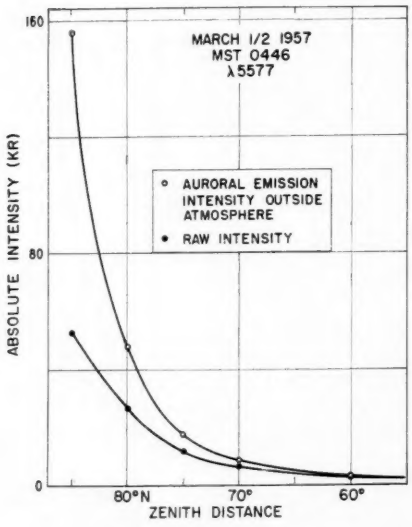


Fig. 3. Auroral intensity of the λ5577 emission before and after corrections for atmospheric extinction and scattering have been applied.

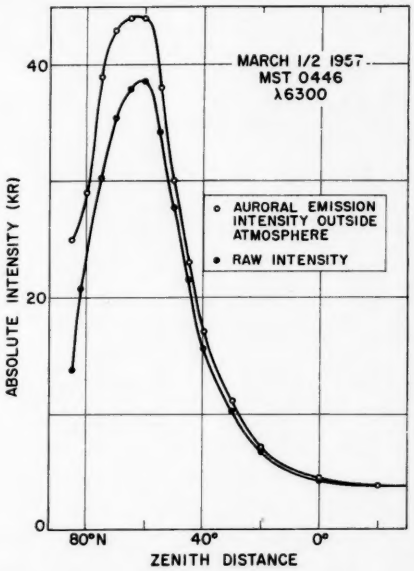


Fig. 4. Auroral intensity of the λ6300–6364 emission before and after corrections for atmospheric extinction and scattering have been applied.

strong wavelength dependence are clearly evident. Since a visual observation of an auroral display corresponds to the uncorrected curve, such observations near the horizon can be very misleading and require careful interpretation. This applies equally well to photographic and spectrographic observations and indicates the desirability of using a sufficiently small field of view to allow the proper corrections to be made as a function of zenith distance. The photometer used in this study incorporated a 5° field.

3. The Photometric Results

3.1. *Intensity*

The observational technique employed yielded both continuous spatial coverage in the meridian and temporal coverage with a 4 min time resolution. Over 600 individual measurements were analysed covering the most interesting and useful time intervals and zenith distances. The complete empirical results are given in tabular form by Rees (1958). The variation of the auroral emissions O*I*(λ5577)

and N.G. $N_2^+(\lambda 3914)$ at $Z = 80°N$ throughout the night is shown in Fig. 5. In addition to the absolute intensity scale as ordinate, the approximate location of the International Brightness Coefficients is indicated in the figure. The diagram illustrates the general morphology of the aurora in these two colours according to the intensity; it does not indicate that the form of the auroral display also varied considerably throughout the night, beginning with a general glow on the horizon and developing into rayed structures around midnight, and gradually fading toward dawn. It may be noted that the $\lambda 5577$ slant intensity at $Z = 80°$ at the beginning of the night, 1 kR, is only slightly higher than the average airglow

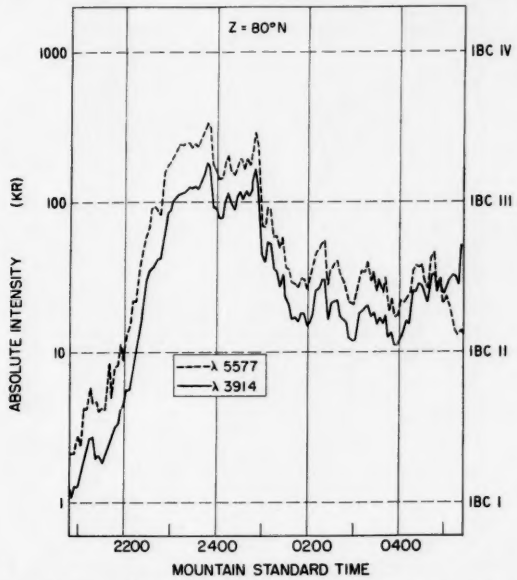


Fig. 5. Variation in the slant intensity of the $\lambda 3914$ and $\lambda 5577$ emissions at a particular zenith distance throughout the night. Data have been corrected for atmospheric extinction and scattered light. The intensity is given in terms of absolute energy units and the visual international brightness coefficient.

intensity of this emission. The close correlation between the intensity of the $\lambda 3914$ and $\lambda 5577$ emissions up to approximately 0400 MST is immediately evident. The reversal of the relative intensity trend after this hour is interpreted in detail as a sunlit aurora in the second paper of this series. Similar results are obtained at the other zenith distances investigated.

If the time parameter is neglected and the two intensities are plotted against each other the correlation shown in Fig. 6 is obtained. We wish to point out especially that this graph includes all the experimental points for zenith distances $Z = 70°$, $75°$ and $80°$ from 2046 MST to 0358 MST inclusive. We conclude that over an intensity range of almost three powers of ten the ratio of the emission intensities is essentially a constant:

$$r = \frac{I(\mathrm{O})5577}{I(\mathrm{O})3914} = 1{\cdot}9 \qquad (3)$$

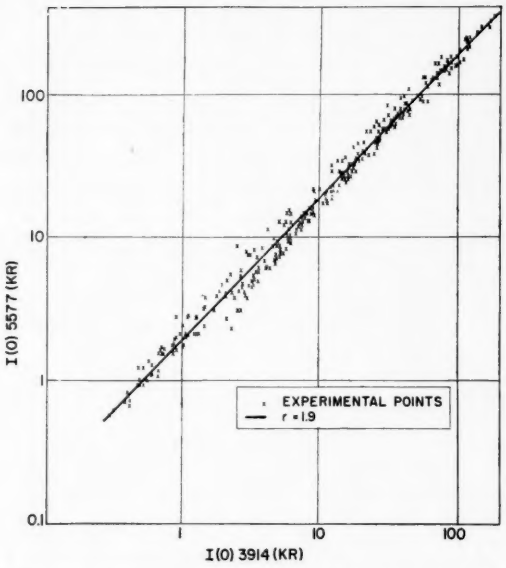


Fig. 6. Correlation between the absolute intensity of λ3914 emission and λ5577 emission. Experimental data include observations at three zenith distances and were obtained during an eight hour period of photometric recording. Data obtained after 0358 MST is not included.

VOL
14
195

Table 1

$$r = \frac{I(\mathrm{O})5577}{I(\mathrm{O})3914}$$

Z Time (MST)	80°	75°	70°	60°	50°	40°	Avg.
2310	2·14	1·54	1·64	2·13	2·37	2·33	2·03
18	2·08	1·85	1·95	2·12	2·20	2·33	2·09
42	1·98	1·64	1·92	1·73	1·73	1·68	1·78
50	1·90	1·61	1·87	1·73	1·78	1·82	1·79
0014	1·73	1·94	2·03	2·20	2·30	2·33	2·09
22	1·70	1·97	1·60	2·02	2·06	2·17	1·92
30	1·67	1·84	1·89	2·07	2·14	2·04	1·94
38	1·60	1·83	1·83	1·94	1·80	1·85	1·81
46	1·75	1·93	1·84	2·08	2·18	2·10	1.98
54	1·71	1·82	1·75	1·91	2·36	2·21	1·96
Avg.	1·83	1·80	1·83	1·99	2·09	2·09	1·94

Because the deflexions on the $\lambda 5577$ graphic record for $Z < 70°$ were, for the most part, too small to be read with accuracy, the ratio at other zenith distances was not included in Fig. 6. However, during the time when auroral intensities were greatest ten sweeps could be found for which readings up to $Z = 40°$ were considered quite reliable. The results of the computation of r at six zenith distances are given in Table 1. The ratio again is essentially constant even though the intensity at the smallest zenith distance, $Z = 40°$, was up to forty times smaller

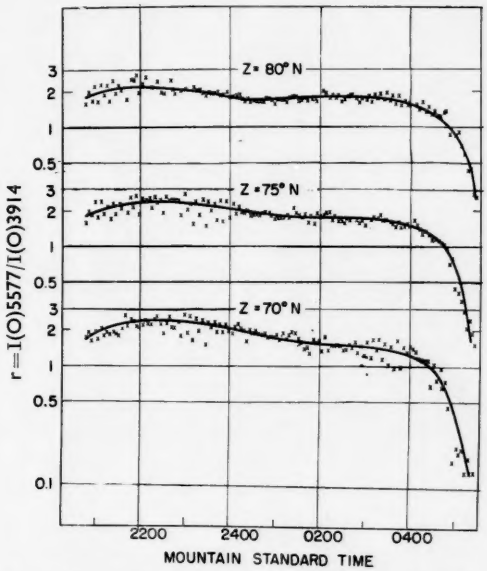


Fig. 7. Variation in the intensity ratio of two auroral emissions as a function of time accounting for much of the scatter shown in Fig. 6.

than at the largest zenith distance considered, $Z = 80°$. The average of the ratios for each sweep and for each zenith distance shows about the same range of values, showing no trend with zenith distance.

Although the ratio, r, is clearly a constant to a first approximation, it would be desirable to account for the scatter in the points in Fig. 6. Plotting r as a function of time the results shown in Fig. 7 are obtained. The scatter of the data noted in Fig. 6 may now be attributed, in part, to a variation with time. Up to about 0400 MST this variation is essentially the same at all three zenith distances. The scatter within each curve is not believed to be of any significance because of the difficulty in making accurate readings of galvanometer deflexions when the curves on the graphic record incorporate a steep gradient (see Fig. 1). This condition prevailed throughout the greater part of the night. The significant enhancement of $\lambda 3914$ during the sunlit aurora is strikingly evident. The slow variation of the ratio, r, as a function of time prior to the onset of the sunlit aurora may be explained as reflecting a variation in the extinction coefficient throughout the night. A constant extinction was assumed in analysing the data; however, the observing log for the night indicated that atmospheric conditions

steadily improved as the night progressed. This would be consistent with higher initial values of extinction and a gradual decrease as a function of time. Results consistent with this hypothesis are obtained at all zenith distances analysed. A variation in r may, however, also arise from a latitude–height effect which will be discussed in the next section.

3.2. *Zenith distance, height and latitude*

A constant ratio of emission intensities λ5577 to λ3914 over a large height interval is highly significant when the excitation mechanisms operative in producing the aurora are considered. A correct interpretation of the data (which have been obtained as a function of zenith distance) with respect to height is therefore

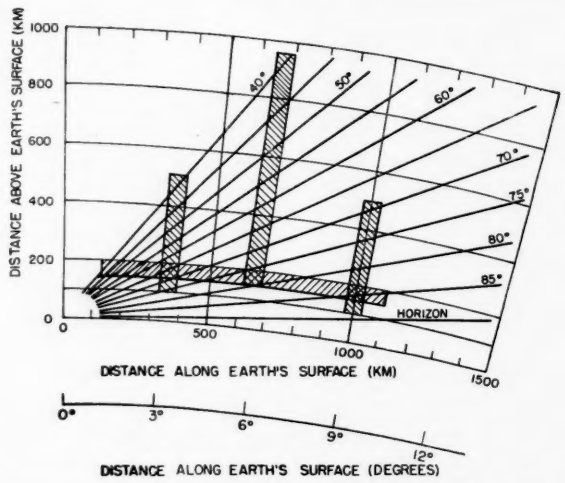


Fig. 8. Meridional cross-section through the earth's atmosphere. Shaded areas indicate possible (idealized) spatial extent of auroral forms illustrating the difficulty encountered in deducing heights from observations obtained as a function of zenith distance.

of utmost importance. Let Fig. 8 represent a cross-section along the meridian sweep, the ordinate representing height above the earth's surface, and the abscissa distance along the earth's surface. The photometer is located at the origin and the radial lines represent lines of sight at various zenith distances. The horizon (at $Z = 90°$) is also shown. In order to explain an emission detectable between $Z = 40°$ and $Z = 85°$ the region of excitation might extend in several possible directions:

(a) The luminosity could all originate from a relatively thin layer located perhaps between 150 and 200 km. If this were the case, reference to Fig. 8 shows that the emitting region would have to extend at least 1000 km, or about 9° in latitude. Moreover, in order to explain the wide variation in intensity with zenith distance, which is many times greater than the variation predicted from a uniformly emitting layer, we must also postulate an intensity gradient as a function of latitude with a strong increase toward the north.

(b) At the other extreme, if the emissive region has a small latitude range

(10 to 20 km) as is usual for auroral forms showing structure such as arcs, rays and draperies, and the same auroral feature is observed between 40° and 85° zenith distance, the required vertical extent of the region would be from 100 km at the lower border to over 1000 km at the upper border. This has been reported in the case of a sunlit aurora, but not for aurorae in the shadow region (STÖRMER, 1955).

(c) Cases (a) and (b) represent two extremes. Most probably, the actual picture lies somewhere in between these extremes. It should be recalled that the results are based primarily on observations at large zenith distances which integrate luminosity from a large region of space. For example, the line of sight at $Z = 80°$ intersects the 100 km altitude layer about 480 km north of the observer, the 150 km level at 650 km, and the 200 km level at 815 km distant (see Fig. 8). Thus, we may be looking at many discrete features at the same time. Just before midnight the observations indicated a rayed aurora and perhaps a drapery. These features are quite thin but very often several parallel curtains can appear simultaneously. The 5° field of view of the photometer adds considerably to the integrating effect.

The aurora of 1–2 March was quite extensive, and it is possible that some form of auroral emission was present everywhere north of Fritz Peak observatory. Unfortunately, this date was prior to the beginning of IGY so that extensive visual observations planned for the IGY were not as yet available. However, a preliminary map compiled by Dr. GARTLEIN's group at Cornell* for this particular night, indicated extensive auroral activity all over the United States; in particular, rayed arcs and rays were reported directly due north (5° and 7° of latitude) of Fritz Peak at various time intervals.

VOL. 14 1959

As a result of the preceding argument a constant intensity ratio $I(O)5577/I(O)3914$ as a function of zenith distance does not necessarily mean a constant ratio with height. In the preceding section a gradual change in r throughout the night (Fig. 7) was interpreted in terms of a change in the extinction coefficient. In view of the discussion just concluded we could also interpret such a change as an aspect effect. A slow drifting of the aurora in latitude would have the effect of presenting to the photometer at a given zenith distance different height layers of the emissive region. A southward drift would result in the lower layers predominating while a northward drift would gradually present the higher layers to the field of view. Thus, if there is a change in the ratio of the intensity r with height it would be expected to appear on the basis of this interpretation. Here also, the integrating effect of observing at large zenith distances would mask any sharp discontinuities.

Although it would be very desirable to obtain luminosity vs. height curves for the aurora, from the present study we can only give luminosity vs. zenith distance curves such as are shown in Fig. 9 for the three emissions recorded and at a particular time of the night. While it might seem, at first, that establishing the lower limit of the auroral height will uniquely determine the height at other zenith distances, the above discussion clearly indicates the necessity of also

* We are grateful to Dr. GARTLEIN for supplying the auroral map compiled for this night.

accounting for the latitude effect, i.e. the "thickness" of the display. Simultaneous observations from several locations would offer a solution to this problem. The results obtained for the O*I* red lines will be discussed in the third paper of this series.

4. Discussion

Although an intensity ratio, r, of approximately two has been reported previously (Vegard and Kvifte, 1945) and the constancy of this ratio for various

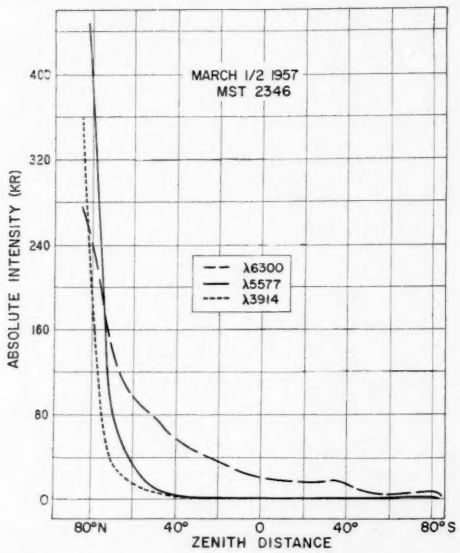


Fig. 9. The auroral emission intensity corrected for atmospheric scattering and extinction is shown as a function of zenith distance of observation.

types of auroral forms was reported by Omholt (1957), the empirical fact that this ratio is more or less a constant over an intensity range of three powers of ten had not been realized before. Such a strong correlation is suggestive of a close relationship in the excitation mechanisms resulting in these auroral features. An explanation may be sought (a) in terms of some simultaneous and proportional excitation mechanisms or (b) a cascading type of excitation mechanism.

In view of the conclusions on the height of the emissions reached in Section (3.2) the question of a constant ratio, r, with height must be left open.

Evidence is accumulating from many sources (Meinel, 1952; Bates, 1955; Seaton, 1954) that the primary mechanism resulting in the N. G. N_2^+ band emission is of the type

$$N_2 + X \rightarrow N_2^{+\prime} + X + e \qquad (4)$$

where X represents the energetic particle and may be a proton, electron, heavy ion or atom. $N_2^{+\prime}$ represents the ionized nitrogen in the excited state. Correlation of auroral phenomena with magnetic activity lead us to consider electrons and protons most seriously. From the experimental evidence associating protons with the N.G. N_2^+ emissions (Meinel, 1952; Meinel and Fan, 1952) the excitation

by protons seems more likely. While a proton stream can account for the λ3914 emission it cannot give rise to the forbidden atomic lines of oxygen, λ5577 and λ6300–6364, since these involve a change in electron spin. This process is one of electron exchange. Simultaneous proportional excitation of the two emission features therefore does not seem probable.

Considering possible cascading mechanisms mention should be made of MITRA's (1948) hypothesis of excitation of N_2^+ by ionization and excitation, followed by

$$N_2^{+\prime} \rightarrow N_2^+ + h\nu \qquad (\lambda 3914) \tag{5}$$

giving rise to the First Negative system, and ionic recombination with negative oxygen ions

$$N_2^+ + O^- \rightarrow N_2 + O(^1S) \tag{6}$$

yielding the upper state of O*I* from which the desired emission is obtained,

$$O(^1S) \rightarrow O(^1D) + h\nu \qquad (\lambda 5577) \tag{7}$$

This set of reactions certainly would dictate a constant ratio and would thus satisfy this aspect of the experimental results. However, from the absolute intensity measurements of the green line and the derived photon emission rate per unit volume (assuming any reasonable thickness for the emission layer) the required equilibrium concentration of O^- exceeds the predicted concentration by several orders of magnitude. Assuming very high electron concentrations would not necessarily alleviate the problem since disappearance of N_2^+ ions by dissociative recombination is probably the more important process accounting for their removal.

The other possible cascading mechanism involves the secondary electrons produced by the primary ionization process of molecular nitrogen, given by equation (4). The process would be excitation of atomic oxygen by electron impact:

$$O(^3P) + e \rightarrow O(^1S) + e \tag{8}$$

This process has already been suggested as a very effective mechanism for producing the forbidden oxygen lines (BATES, 1955; SEATON, 1954). The rate coefficient for the reaction and the oxygen abundance are sufficiently high to account easily for the observed emission intensities. The constant ratio of the emission intensities λ5577 to λ3914 over a thousandfold range of intensity, throughout the entire aurora in the dark atmosphere, and for various types of auroral forms implies that (a) excitation by electron collision is by far the dominant mechanism, and (b) that these electrons originate from the ionization excitation of N_2, i.e. it is this process which gives rise to the active electrons capable of producing the O*I* green line and probably other emissions resulting from electron collision. The oxygen red lines, to be discussed in detail in paper III, do not show a one to one correlation, according to intensity, with the green line, rather the ratio of intensities appears to be a function of the intensity itself.

A quantitative prediction of the ratio $I(O)5577/I(O)3914$ would require knowledge of several factors. The relative population of the various electronic levels of N_2^+ following the primary collisional ionization of N_2 must be known if

VOL. 14 1959

the ratio of the number of 3914 quanta to the number of secondary electrons emitted is to be determined. Ionization cross-sections of N_2 by *proton* impact are not available at present. Massey and Burhop (1952) give the apparent cross-section for ionization of N_2 by *electron* impact and we shall use these values to illustrate the computation. For 100 eV electrons the cross-section is $2{\cdot}87 \times 10^{-16}$ cm^2. Stewart (1956) has measured the probability of excitation of the (O, O) Negative band of N_2^+. From the published curves a cross-section of $6{\cdot}2 \times 10^{-18}$ cm^2 for 100 eV electrons is obtained. Therefore, $2{\cdot}87 \times 10^{-16}/6{\cdot}2 \times 10^{-18} = 46$ ionizations occur per λ3914 photon emitted, producing that number of secondary electrons. Some of these electrons will be sufficiently energetic to produce one or more additional ionizations. The final energy distribution of the secondary and tertiary electrons must be determined. Recently Bates, McDowell and Omholt (1957) have calculated the energy distribution of secondary electrons ejected from the outer shell of neon by proton impact. If one may apply the results qualitatively to nitrogen it appears that the energy distribution can be sharply peaked in a small energy range. This might well account for the selective excitation of the ^1S state of O*I*. A quantitative prediction of the ratio, r, cannot be made at present, however, there seem to be a sufficient number of secondary electrons to account for a ratio, $r = 1{\cdot}9$.

5. Summary

(a) The absolute intensities of the N.G. $N_2^+(\lambda 3914)$ emission and the O*I*(λ5577) line were determined for one low latitude aurora.

(b) The ratio of the intensities, $r = I(\mathrm{O})5577/I(\mathrm{O})3914$, was found to be a constant over an intensity range of three powers of ten and throughout the entire night regardless of the auroral type, but for emission from the dark atmosphere only. The numerical value for this ratio was found to be $r = 1{\cdot}9$. Considering the uncertainties in the absolute calibration of each telescope the value given is probably correct to within ± 50 per cent. An error in the absolute calibration would enter systematically in the numerical value of this ratio but would not affect its constancy. Variations from the average ratio were interpreted as arising either from a change in the extinction coefficients throughout the night, or as a height–latitude effect.

(c) The excitation of O*I*(λ5577) can occur primarily by collision with the secondary electrons resulting from the primary ionization of N_2, yielding N_2^+. The requirement of secondary electrons to explain the empirical correlation favours the simultaneous ionization and excitation of N_2 over the excitation of N_2^+ from its ground state to one of the excited states.

Acknowledgements—It is a pleasure to acknowledge the help of Mr. E. Marovich who shared in the arduous task of taking continuous observations throughout a long night. I am grateful to Dr. F. E. Roach for many stimulating conversations during the course of this study.

References

Ashburn E. V. 1953 *J. Atmosph. Terr. Phys.* **5**, 83.
Barbier D. and Pettit H. 1952 *Ann. Géophys.* **8**, 232.

VOL
14
195

Bates D. R.	1955	*Ann. Géophys.* **11,** 253.
Bates D. R., McDowell M. R. and Omholt A.	1957	*J. Atmosph. Terr. Phys.* **10,** 52.
Chamberlain J. W. and Oliver N. J.	1953	*Astrophys. J.* **118,** 197.
Chandrasekhar S.	1950	*Radiative Transfer.* Oxford University Press.
Hunten D. M., Roach F. E. and Chamberlain J. W.	1956	*J. Atmosph. Terr. Phys.* **8,** 345.
Marlow D. G. and Pemberton J. C.	1949	*Rev. Sci. Instrum.* **20,** 724.
Meinel A. B.	1952	*Mem. Soc. Roy. Liege* **12,** Fsc. I–II, 203.
Meinel A. B.	1953	*Astrophys. J.* **118,** 200.
Meinel A. B. and Fan C. Y.	1952	*Astrophys. J.* **115,** 330.
Mitra S. K.	1948	*The Upper Atmosphere.* Royal Asiatic Society, Bengal, Calcutta.
Omholt A.	1957	*Astrophys. J.* **126,** 461.
Rees M. H.	1958	Ph.D. Thesis, University of Colorado.
Roach F. E. and Pettit H.	1951	*J. Geophys. Res.* **56,** 325.
Seaton M. J.	1954	*J. Atmosph. Terr. Phys.* **4,** 295.
Stewart D. T.	1956	*Proc. Phys. Soc. Lond.* A **69,** 437.
Störmer C.	1955	*The Polar Aurora.* Oxford University Press.
Vegard L. and Kvifte G.	1945	*Geofys. Publ.* **16,** No. 7.

VOL.
14
1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 338 to 344. Pergamon Press Ltd. Printed in Northern Ireland

Absolute photometry of the aurora—II

Molecular nitrogen emission in the sunlit atmosphere

M. H. REES*
Boulder Laboratories, National Bureau of Standards and University of Colorado,
Boulder, Colorado

(*Received* 30 *August* 1958)

Abstract—The absolute emission intensity of a sunlit aurora was determined for the 1.N.G. λ3914 band of N_2^+ by subtracting the emission intensity originating in the dark atmosphere from the total measured intensity. The method is based on the constancy of the ratio of emission intensity of the oxygen green line O*I* λ5577 to the λ3914 emission in the dark atmosphere. The ion density of N_2^+ is computed as a function of height from the emission rate per unit volume.

1. INTRODUCTION

IDENTIFICATION of sunlit aurorae in the past have been based primarily on photographs taken simultaneously from several locations from which the height of an identifiable emission feature was determined. If the height was established to be in the sunlit atmosphere it was concluded that the display constituted a sunlit aurora. Occasionally divided rays have been observed, the lower portion lying in the dark atmosphere, separated from the upper segment in the sunlit atmosphere by a region of greatly reduced luminosity. Sunlit aurorae have been discussed extensively by STÖRMER (1955).

An observer or instrument on the ground will measure the total luminosity in a column extending from the instrument to the upper limit of the emission height at the given zenith distance. One cannot, *a priori*, attribute the total intensity to a sunlit aurora since there is no basis for assuming that emission in the dark atmosphere ceases during the hours after sunset and before sunrise. At large zenith distances, especially, the line of sight traverses a considerable portion of the region from which dark atmosphere auroral emission may originate. A more sensitive procedure for isolating a sunlit aurora is clearly necessary if a quantitative analysis of the photon emission rate is to be carried out.

Among the auroral emissions enhanced in the sunlit atmosphere the λ3914 band of the first negative system of N_2^+ is one of the most prominent. The approach used in this paper for studying the sunlit aurora is based on the photometric enhancement of λ3914 around dawn. Photometric measures of three auroral emissions isolated by means of narrow band interference filters were carried out throughout the entire night of 1–2 March 1957. These were the N_2^+ λ3914 band and the O*I* λ5577 and O*I* λ6300–6364 lines. In paper I of this series (REES, 1958) a nearly constant ratio between the intensity of O*I* (λ5577) and N_2^+ (λ3914) was shown to exist for emission from the dark atmosphere. We assume (a) that there is no significant enhancement of λ5577 in the sunlit atmosphere, and (b)

* Now at the Geophysical Institute, University of Alaska.

that the ratio $r = I(\mathrm{O})5577/I(\mathrm{O})3914$ for dark atmosphere emission remains constant into dawn. The component of the λ3914 intensity originating from the dark atmosphere may therefore be computed; subtraction from the total intensity yields the λ3914 emission from the sunlit atmosphere.

2. The Observations

The morphology of the auroral intensity at a zenith distance of 80° throughout the night was shown in Fig. 5 of Paper I. We noted the close correlation between

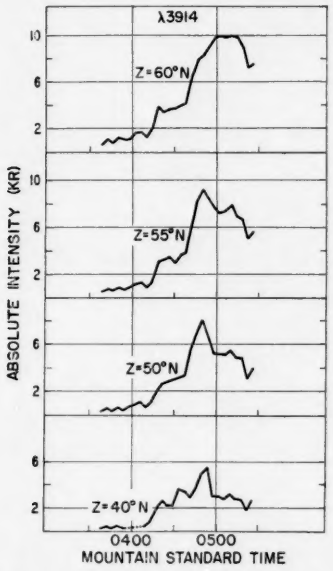


Fig. 1. Enhancement of the auroral emission λ3914 toward dawn. The absolute intensity has been corrected for atmospheric scattering and extinction.

the two emission intensities λ5577 and λ3914 and the departure from this correlation in the hours just before sunrise. This was illustrated more definitively in Fig. 7 of Paper I from which the extent of the enhancement of λ3914 over λ5577 is clearly evident. For example, at $Z = 70°$ N at Mountain Standard Time 0526 the λ3914 intensity amounted to almost nine times the intensity of the λ5577 emission as compared to the average ratio for the night (up to 0358 MST) $r = I(\mathrm{O})5577/I(\mathrm{O})3914$, of 1·9. The increase in intensity attributed to the sunlit aurora occurred not only at the three zenith distances generally used in this study, $Z = 70°, 75°, 80°$ N, but also at other zenith distances for which readable deflexions were available. Typical curves are shown in Fig. 1 of the present paper. These curves represent the total λ3914 intensity, not the sunlit aurora alone.

3. The Analysis

The general method of analysing the observations, including correction for extinction and the scattered light, leading to results in terms of absolute intensity units is given in Paper I.

In Fig. 2 the height, h_0, of the shadow line is to be determined as a function of time and zenith distance of observation, Z. To allow for Rayleigh scattering in the lower atmosphere and absorption due to ozone a screening height, h_s, of 25 km

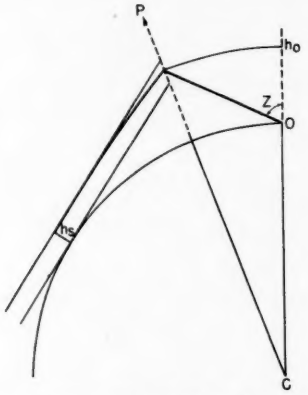


Fig. 2. Geometry of the sunlit aurora. The line of sight from an observer at O at a zenith distance Z intersects the sun's rays at a height h_0. h_s is the screening height and the curvature of the ray represents atmospheric refraction.

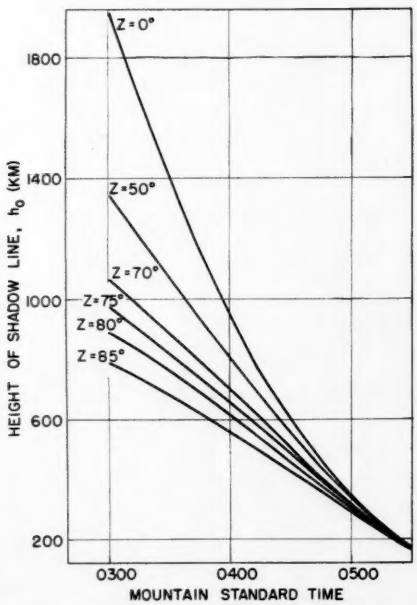


Fig. 3. The height of the shadow line at Fritz Peak on 2 March 1957 as a function of time. The curves given are for observations due north at various zenith distances.

is adopted. A correction for refraction in the atmosphere is included. Details of the problem are given by CHAMBERLAIN (1958). The time parameter is, of course, determined by the position of the sun; thus the results apply to a particular day only (as well as a particular latitude) although the equations are quite general.

The height calculations were carried out on an electronic computing machine. The results for the morning of 2 March 1957 are shown in Fig. 3 for several zenith distances of observation.

The line of sight at a given zenith distance traverses the portion of the atmosphere in the shadow first, then passes into the sunlit portion. Any emission originating in the shadow portion must therefore be subtracted from the total measured intensity to isolate the emission from the sunlit atmosphere. The general constancy of r throughout the night with only a very small and gradual

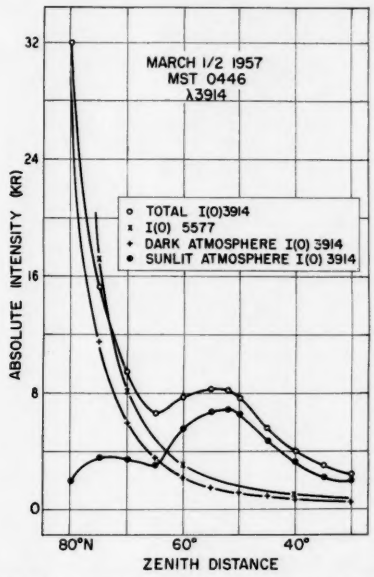


Fig. 4. Steps in the analysis of the sunlit aurora given in the text.

change as a function of time up to about 0330 MST leads to the assumption that, as far as the emission from the shadow region is concerned, no drastic change in this ratio is to be expected. The λ3914 emission from the shadow region may, therefore, be computed by dividing I(O)5577 by the appropriate ratio. It is also assumed that there is no enhancement of λ5577 in the sunlit atmosphere. Although, in principle, such an enhancement could occur, requiring photons of wavelength less than 2972 Å, the phenomenon has not been reported so far. Moreover, the experimental data do not extend below a height of 200 km where the oxygen abundance already is much less than the maximum at about 100 km. From the curves in Fig. 7 of Paper I the ratio, r, is obtained by extrapolation, $r = 1{\cdot}6$ for $Z = 80°$, $r = 1{\cdot}5$ for $Z = 75°$, and $r = 1{\cdot}4$ for $Z = 70°$. The steps in the computation of I(O)3914 sunlit are illustrated graphically in Fig. 4 for a portion of the sweep at 0446 MST. A factor $r = 1{\cdot}4$ was used in the reduction of this data for zenith distances of 70° and less.

The photometric hump between $Z = 60°$N and $Z = 40°$N appeared only on a few sweeps and is not representative of the general intensity distribution in the sunlit aurora. It does, however, illustrate the significant difference between the

VOL. 14 1959

λ5577 and λ3914 emissions during this portion of the auroral display. Using all the experimental points for $Z = 70°$, 75° and 80° N from 0330 to 0526 MST the sunlit λ3914 intensity was computed. From the curves in Fig. 3 the height of the shadow line at various times is determined. The combined data for the three zenith distances yield the curve in Fig. 5. In spite of the scatter of the experimental points the data indicate that the sunlit aurora first becomes measurable at a height just over 700 km.

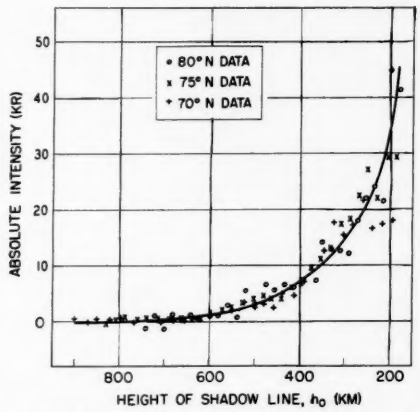


Fig. 5. Absolute intensity of the sunlit aurora for the λ3914 emission. The intensity at a given height, h_0, refers to the total slant intensity from a column extending upwards from h_0.

It should be recalled that an observation at a zenith distance of 70° extending between a height of 200 km and 700 km covers a latitudinal range of about 850 km, while for $Z = 80°$ the corresponding range is approximately 1500 km. Since the height of the shadow line is a function of latitude and since there is no reason to assume the cause of the excitation to be uniform over such a wide range of latitude, the agreement in the height determination for the three zenith distances considered strengthens the validity of our result. Although the data at smaller zenith distances could not be treated quantitatively the curves in Fig. 1 show agreement with the hypothesis that the increase in intensity toward dawn is caused by a sunlit effect.

4. Discussion

Two possible excitation mechanisms present themselves to explain the 1.N.G. N_2^+ emissions from an atmospheric region illuminated by sunlight. There is simultaneous photo-ionization and excitation to the required electronic level:

$$N_2(X^1\Sigma_g^+) + h\nu\ (<660\ \text{Å}) \rightarrow N_2^+(B^2\Sigma_u^+) + e \qquad (1)$$

or, "sensitization" of the atmosphere by incoming corpuscles producing N_2^+ ions along their track followed by resonance scattering:

$$N_2^+(X^2\Sigma_g^+) + h\nu\ (<3914\ \text{Å}) \rightarrow N_2^+(B^2\Sigma_u^+) \qquad (2)$$

Although the normal energy distribution in the solar flux provides only a very small number of photons with an energy greater than 18·74 eV this may not be

the case during an aurora. Soft X-rays (presumably produced as bremsstrahlung from incoming electrons) have been detected from balloon-borne instruments at approximately the 30 km level simultaneously with visual observations of auroral displays.* However, aside from this consideration, BATES (1949b) has shown that resonance scattering is by far the more important process unless the electron density is unusually high. Thus, a high degree of "sensitization" of the atmosphere is required. For resonance excitation the photon emission is:

$$\pi_r = P_{3914}[N_2^+] \tag{3}$$

where P_{3914} is the $\lambda 3914$ photon emission probability per illuminated ion per sec and $[N_2^+]$ represents the number of N_2^+ ions per cm³. P has been calculated

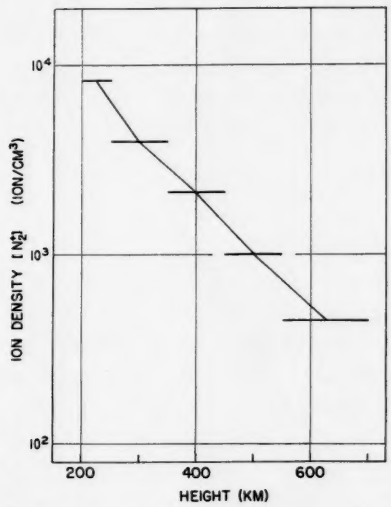


Fig. 6. The variation of the N_2^+ ion density as a function of height in a sunlit aurora.

theoretically by BATES (1949a) and slightly revised by SHULL (1951), its value being $P = 0{\cdot}3$ sec^{-1}. Using the data from the smoothed curve in Fig. 5 the N_2^+ density may be calculated at various heights. Recalling that the intensity given for each height refers to a column extending from the height indicated to the upper limit from which emission is detectable, 700 km, the data is first reduced to intensity values referred to certain height intervals. An observation of the 600 km height layer made at a zenith distance of 75° is about 1500 km distant from the observing site. With a field of view of the photometer of 5° the spread in height is therefore 130 km. This height resolution becomes better at the lower altitudes. The height interval chosen, $h_1 - h_2$, is governed by the above considerations.

Assuming isotropic emission and no self-absorption the emission rate per cm³ (π_r) may be obtained.

$$\frac{\Delta I(\mathrm{O})\ 3914\ 10^9 \text{quanta cm}^{-2}\,\text{sec}^{-1}\,\text{col}^{-1}}{h_1 - h_2\ (\text{cm-col}^{-1})} = \pi_{(3914)}\ (\text{quanta-cm}^{-3}\,\text{sec}^{-1}) \tag{4}$$

* Reported in the IGY Bulletin No. 10, p. 9 (April 1958) of the National Academy of Sciences.

VOL. 14 1959

Since the curve in Fig. 5 represents a composite of data obtained at several zenith distances a correction reducing slant intensity to vertical intensity need not be made. Taking resonance scattering as the excitation process the N_2^+ density may be computed from equation (3). The calculations are given in Table 1 and the results are shown in Fig. 6.

Table 1

h_0 (km)	I(O)3914 (kr)	$h_2 - h_1$ (km)	ΔI(O)3914 (kr)	π_{3914} (Q-cm^{-3}-sec^{-1})	$[N_2^+]$ (ion-cm^{-3})
650	0·2				
600	1·0	550–700	2·0	133	446
550	2·0				
500	3·4	450–550	3·0	300	1000
450	5·0				
400	7·4	350–450	6·4	640	2136
350	11·4				
300	16·6	250–350	12·0	1200	4000
250	23·4				
200	36·0	200–250	12·6	2520	8400

BATES (1949a) has estimated the N_2^+ concentration from twilight enhancement of the λ3914 emission, arriving at a figure several orders of magnitude smaller than we have found at the height of the F-region. The higher concentration found during the aurora is, possibly, due to a corpuscular stream "sensitizing" the atmospheric constituents.

Acknowledgement—I am indebted to Mr. L. R. MEGILL for programming for use on an electronic computing machine the height calculation leading to Fig. 3.

REFERENCES

BATES D. R.	1949a	*Proc. Roy. Soc.* A **196,** 562.
BATES D. R.	1949b	*Proc. Roy. Soc.* A **196,** 217.
CHAMBERLAIN J. W.	1958	*Astrophys. J.* **127,** 54.
REES M. H.	1958	*J. Atmosph. Terr. Phys.* In press.
SHULL H.	1951	*Astrophys. J.* **114,** 546.
STÖRMER C.	1955	*The Polar Aurora.* Oxford University Press.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 345 to 347. Pergamon Press Ltd. Printed in Northern Ireland

Spiral occurrence of sporadic-*E*

E. L. Hagg, D. Muldrew and E. Warren
Defence Research Board, Shirley Bay, Ottawa

(*Received* 11 *August* 1958)

A study of high density sporadic-*E* ionization at high latitudes has revealed that it is distributed in the form of a spiral similar to that of Störmer's precipitation curve for negative particles (Störmer, 1917). The phenomenon tabulated as sporadic-*E* may have a variety of causes, among which are thunderstorms, turbulence in a reflecting region, ionized meteor trails and the precipitation of particles from the sun (Smith, 1945a). In an attempt to isolate sporadic-*E* of a single variety, only those events were considered in which vertical incidence reflection of a frequency greater than 5 Mc/s occurred. For some stations at which intense sporadic-*E* is frequent (see Table 1) a minimum frequency of 7 Mc/s was chosen.

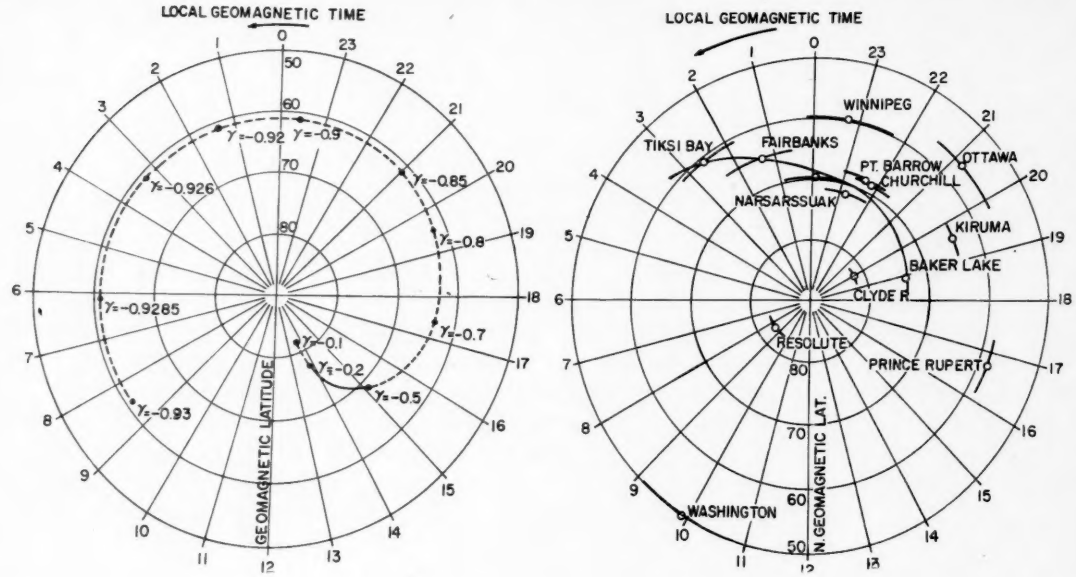
Table 1

Station	Geomagnetic latitude	Lower limit of fE_s Mc/s	Years' data used
Resolute	82·8°	5	1949–1954
Clyde River	81·9°	5	1944–1950
Baker Lake	73·8°	7	1950–1954
Narsarssuak	71·6°	5	1951–1954
Ft. Chimo	69·7°	7	1949–1953
Churchill	68·6°	7	1951–1955
Pt. Barrow	68·2°	7	1950–1954
Fairbanks	65·1°	5	1948–1954
Kiruna	64·4°	5	1949–1954
Tiksi Bay	60·8°	5	1950–1952
Winnipeg	59·7°	5	1950–1954
Prince Rupert	58·5°	5	1949–1953
Ottawa	56·8°	5	1950–1956
Washington	50·0°	5	1948–1954

The data from Narsarssuak, Kiruna, Fairbanks and Washington were obtained from a National Bureau of Standards report (Smith, 1945b); those for Tiksi Bay were obtained from an AVCO report (Coroniti and Penndorf, 1958). The remainder of the data was obtained for the Defence Research Board by the Department of Transport.

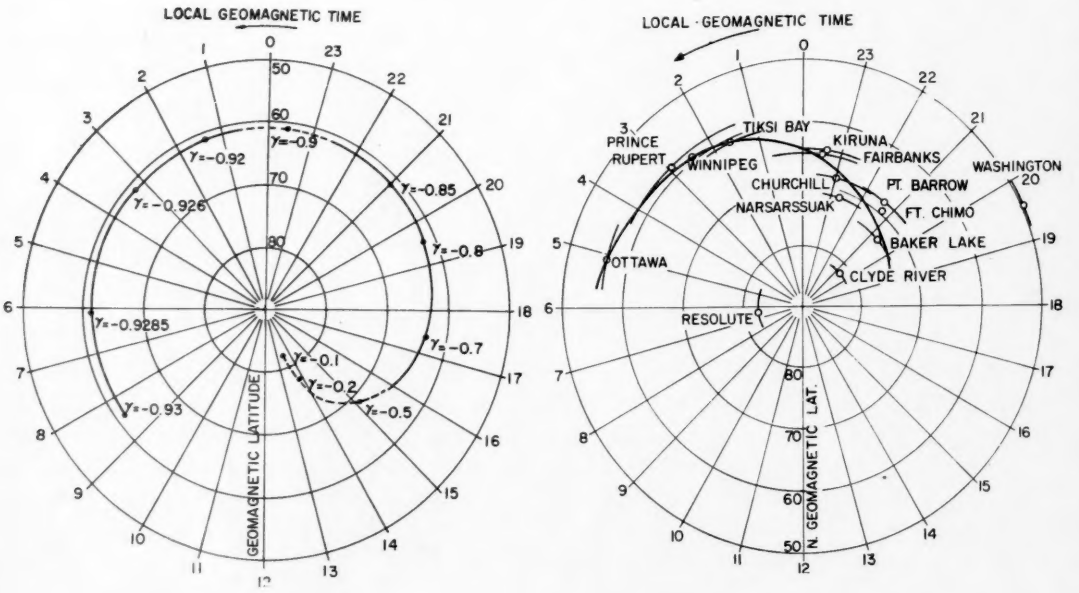
The diurnal variation of the total number of occurrences of fE_s greater than the stated minimum frequency, observed during the month of June for several years, was plotted for each station. The process was then repeated for the month of December. The time of day for which sporadic-*E* was most frequent was

VOL. 14 1959



(a) (b)

Fig. 1. (June): (a) Störmer's precipitation spiral; (b) distribution of high density sporadic-E ionization.



(a) (b)

Fig. 2. (December): (a) Störmer's precipitation spiral; (b) distribution of high density sporadic-E ionization.

determined. Also found was the fraction of the day during which the number of occurrences was greater than 95 per cent of the maximum value. These times and intervals were expressed in local geomagnetic time (McNish, 1936) and were plotted against the geomagnetic latitude of the stations, as shown in Figs. 1(b) and 2(b).

During December the spiral is well defined between 55° and 75° N geomagnetic latitude; in June, however, the southern limit increases to about 60° N geomagnetic latitude. In June the maximum for Kiruna does not fit the spiral, but a secondary maximum for Kiruna does.

Störmer's spiral for negative particles precipitating in the northern latitudes is shown in Figs. 1(a) and 2(a). His theory predicts a ring current flowing around the earth parallel to the geomagnetic equatorial plane. The latitude at which the circular section of the precipitation spiral ($\gamma = -0{\cdot}80$ to $\gamma = -0{\cdot}93$) occurs is dependent on the momentum of the particles and upon the magnetic field produced by the ring current. The spirals of Figs. 1(a) and 2(a) have been drawn on the assumption that the ring current and the momentum of the particles are adequate to cause the circular section of the spiral to coincide with the observed latitude of the auroral zone. The integration constant γ is one of the parameters that describe the trajectory of a charged particle in the field of a magnetic dipole. The range of values applicable to trajectories between the sun and the earth depends on the position of the sun in the geomagnetic co-ordinates of the earth. The solid lines of Figs. 1(a) and 2(a) show the range of γ-values possible for June and for December.

The theoretical spirals and the observed spirals have only approximately the same form. Störmer, however, derived his spirals from a relatively simple model and on the assumption of a constant ring current; detailed agreement is therefore not to be expected. Under the circumstances, the fact that spirals are observed is support for Störmer's ideas on the nature and origin of the aurora.

VOL. 14 1959

References

Coroniti S. C. and Penndorf R. 1958 *Propagation Conditions in the Arctic* p. 60. *Scientific Report No.* 1 issued by AVCO Manufacturing Corporation, Boston, Massachusetts.

McNish A. G. 1936 *Terr. Magn. Atmos. Elect.* **41,** 37.

Smith E. K., Jr. 1955a *Circular* 582 p. 12. U.S. Department of Commerce National Bureau of Standards, Central Radio Propagation Laboratory, Boulder, Colorado.

Smith E. K., Jr. 1955b *Circular* 582 p. 137. U.S. Department of Commerce, National Bureau of Standards, Central Radio Propagation Laboratory, Boulder, Colorado.

Störmer C. 1917 *Terr. Magn. Atmos. Elect.* **22,** 23 and 97.

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 348 to 349. Pergamon Press Ltd. Printed in Northern Ireland

RESEARCH NOTE

An additional lunar influence on equatorial E_s at Huancayo

(*Received* 5 *February* 1959)

RECENTLY a lunar effect in the times of sudden day-time disappearance of equatorial E_s has been demonstrated (MATSUSHITA, 1957). The purpose of this note is to show an additional influence of the moon on this type of sporadic-E at Huancayo, Peru (geomagnetic latitude S 0·6°, magnetic dip +2°).

Equatorial E_s has been studied in some detail by RAWER (1955), MATSUSHITA (1951, 1953), SMITH (1957) and others. An example of E_s-q (international designation for equatorial E_s) at Huancayo is shown on the ionogram reproduced as Fig. 1. The characteristics of E_s-q are given below.

(a) It regularly occurs during the daylight hours in a narrow belt along the magnetic equator.
(b) It is always partially transparent to probing radio waves; that is, it never blankets radio reflections from higher layers.
(c) It usually shows a well-defined lower edge lying between 100–110 km giving scattered and diffuse echoes above the principal echo.
(d) In well-developed cases, the diffuse echoes are contained below a sharp upper boundary that starts at about f_0E and increases in height with increasing frequency.
(e) The maximum frequency returned, (f_0E_s), usually ranges from 8 to 13 Mc/s primarily depending on the proximity of the ionosonde to the magnetic equator.
(f) Multiple echoes are not observed.

In the period studied (May 1957–April 1958), E_s-q was observed every day at Huancayo. It began as early as 0600 on some days and as late as 1000 on others. On some days it would disappear in the early afternoon for periods of 15 min to several hours, the effect studied by MATSUSHITA, but the final late afternoon disappearance occurred regularly between 1730–1830 75°W time. The variability of the time of first appearance of E_s-q in the morning suggested further study. The 12 month period was divided into the three seasons and the time of the first appearance of E_s-q at Huancayo on each day was tabulated. (For the purpose of this analysis, the time tabulated was the time at which E_s-q first appeared and continued to be observed on at least the succeeding three 15 min ionograms. Prior isolated occurrences were ignored.) The days were then grouped by approximate lunar age and the mean time of first appearance computed. The mean points (X's), along with the 3 day running mean to smooth out spurious fluctuations, are shown for the three seasons in Fig. 2. A lunar dependence is clearly visible in the equinox and December solstice curves, with E_s-q tending to occur earlier during times of new and full moon (lunar age 0 and about 14 days). A considerably weaker lunar effect of questionable significance appears to be present in the June solstice data. In his earlier work on the sudden mid-day disappearance of E_s-q, MATSUSHITA assumed that E_s-q was caused by vertical upward drift forces due to eastward flowing electric currents and he supposed that westward flowing currents caused by lunar tidal forces were responsible for weakening the E_s-q and, thus, the lunar effect he observed. From the work of CHAPMAN and BARTELS (1940) it can

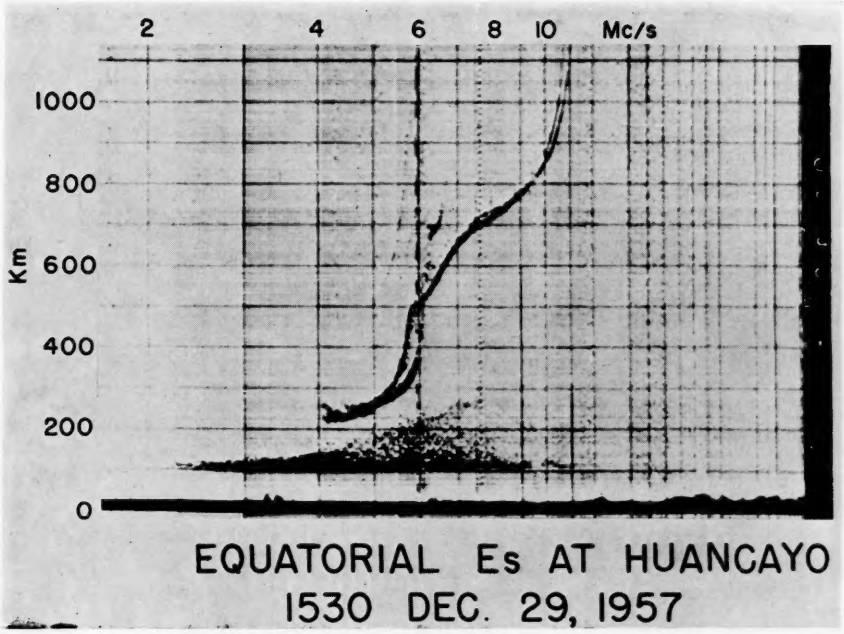


Fig. 1. Time of first appearance of E_s-q at Huancayo (3 day running mean plotted).

VOL
14
195

be seen that during new and full moon the contribution of the current system attributed to the lunar tidal force is such that it *adds* to eastward flowing solar-associated current (S_q) during the early morning hours (0600–1000) whereas during the first and third quarter the lunar contribution *weakens* the eastward flowing current during the early morning

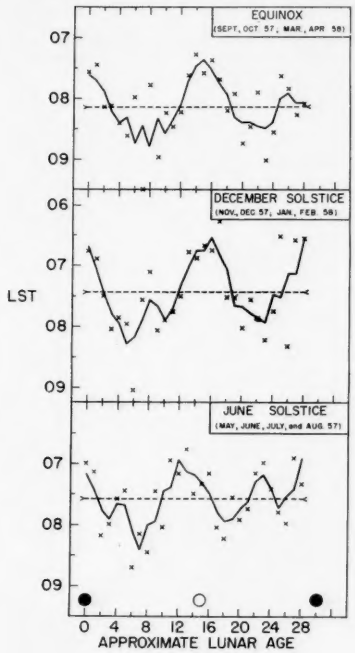


Fig. 2

hours. Thus, if E_s-q depends on the strength of the eastward current, it should appear earlier during new and full moon and later at the first and third quarters, which is exactly the kind of dependence shown above. Therefore, MATSUSHITA's hypothesis on the relation between the electric current system and E_s-q seems to have additional support.

The observations reported here were taken by personnel of the Huancayo Geophysical Observatory. By dependably providing high quality ionograms and careful reductions of the data, they have contributed substantially to this study.

R. W. KNECHT

National Bureau of Standards
Boulder, Colorado

REFERENCES

MATSUSHITA S.	1957	*J. Atmosph. Terr. Phys.* **10,** 163.
RAWER K.	1955	*C. R. Acad. Sci. Paris,* **237,** 1102.
MATSUSHITA S.	1951	*J. Geomagn. Geoelect.* **3,** 44.
MATSUSHITA S.	1953	*J. Geomagn. Geoelect.* **5,** 109.
MATSUSHITA S.	1953	*Rep. Ionos. Res. Japan* **7,** 45.
SMITH E. K.	1957	*Nat. Bur. Stand. Circular* **582,** 37.
CHAPMAN S. and BARTELS J.	1940	*Geomagnetism* Vol I, p. 267, Fig. 11. Oxford University Press.

VOL. 14 1959

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, p. 350. Pergamon Press Ltd. Printed in Northern Ireland

Book review

Handbuch der Physik, Band 50. **Astrophysik I.** Sternoberflächen—Doppelsterne. Edited by S. FLÜGGE. Springer Verlag, Berlin, 1958. 458 pp., DM98.

IT is well over 20 years ago since the Springer Verlag published its six-volume *Handbuch der Astrophysik* which, in spite of some unevenness, gave a remarkably complete picture of the observational data of the time. It provided also an early and most valuable introduction into the then still young field of theoretical astrophysics. However, the *Handbuch* appeared at a time of rapid progress, and a supplementary volume, published in 1936, could hardly catch up with new developments.

In the meantime, improved instrumentation and new techniques such as that of radio-astronomy have widened the observational basis of astrophysics at an ever increasing rate, and the application of nuclear physics to stellar structure and stellar evolution has profoundly influenced the general trend of astrophysical research so that the old *Handbuch* has for quite some time been sadly out of date.

With the present volume, which is to be followed by four others, we witness its resurrection in a new form. Being part of the new *Handbuch der Physik* the astrophysics volumes are now addressed primarily to physicists, amongst whom, after all, present-day astronomers form only a special group. The first three of the five volumes are to deal with the physics of the observable outer layers of the stars, with the theory of their internal structure, and with the particular study of the Sun and of the Solar system. The fourth and fifth volumes are to be devoted to the composition of our own and other galaxies, and to astrophysical instrumentation.

The present first volume is made up of ten separate contributions only two of which, and short ones at that, are in German. More than half of the volume is in French, the remainder in English.

The first five articles are devoted to basic data of stellar spectroscopy. FEHRENBACH gives on some 90 pages an extremely useful account of the classification of stellar spectra, of the various systems of classification with one or two parameters and of the physical basis of such and possible other parameters. His discussion of the very important stars with "peculiar" spectra is expanded in a special article by KEENAN who deals with the question of differences of chemical composition. SWINGS contributes a 30-page résumé on the role of molecular spectra in astrophysics. A wealth of entirely new observational data is contained in the article by GREENSTEIN who deals with the spectra of white dwarf stars now available to us thanks to the 200 in. telescope. Two contributions by WURM deal with observation and theory of planetary nebulae.

The theoretical interpretation of stellar spectra forms the contents of an article by BARBIER who manages to give on some 120 pages a concise and most lucid account of the present state of the theory of stellar atmospheres.

The three remaining articles, placed between the observational and theoretical ones, are devoted to visual, eclipsing and spectroscopic double stars. The first, by VAN DE KAMP, on orbits and masses of visual binaries is the most "astronomical" article in the volume. GAPOSCHKIN discusses amongst others the determination of stellar dimensions, and STRUVE and SU-SHU-HUANG deal with the particular phenomena such as gaseous motions which appear in a study of close spectroscopic binaries.

It remains to say that the volume contains a fairly extensive general index in English and German, and a shorter one in French which refers to the French articles only. As one would expect, the volume is beautifully produced. It is impossible, of course, to reproduce in a book all the detail of original spectrograms, but FEHRENBACH's spectra in particular show up remarkably well. The volume will become an essential companion of every astrophysicist; it should be of much interest to spectroscopists in general.

H. A. BRÜCK

VOL
14
195

Journal of Atmospheric and Terrestrial Physics, 1959, Vol. 14, pp. 351 to 352. Pergamon Press Ltd. Printed in Northern Ireland

Papers to be published in future issues

P. C. Clemmow and M. A. Johnson: A contribution to the theory of the motion of weak irregularities in the ionosphere

G. F. Walton: Calculation of the vertical distribution of atmospheric ozone

N. F. Barber and D. D. Crombie: VLF reflections from the ionosphere in the presence of a transverse magnetic field

L. G. H. Huxley: A discussion of the motion in nitrogen of free electrons with small energies, with reference to the ionosphere

J. M. Malville: Type-*B* aurorae in the Antarctic

W. Becker: New methods and some results concerning true ionospheric height calculation

P. Bandyopadhyay: The early morning *E2*-layer and some evidence of pre-sunrise *F*-layer splitting

W. L. Price: Anomalous ionospheric reflection during solar eclipses

J. D. Whitehead: Sudden changes in the virtual height of radio waves reflected from the *E*-region of the ionosphere

M. D. Vickers: The effect of the *F1*-layer on the calculation of the height of the *F2*-layer

F. Mariana: A new theoretical model of the composite *F*-layer

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—I. Solar variations

C. A. Onwumechilli and N. S. Alexander: Variations in the geomagnetic field at Ibadan, Nigeria—II. Lunar and luni-solar variations in *H* and *Z*

E. Woyk: The refraction of radio waves by a spherical ionized layer

A. J. Ferraro and J. J. Gibbons: Polarization by means of the multislab approximation

B. N. Singh and R. L. Ram: Rhythmic fading of short-wave radio signals

A. G. McNamara and D. W. R. McKinley: The effect of trail irregularities on the interpretation of meteor echoes

K. C. Clarke and A. E. Belon: Spectroscopic observations of the great aurora of 10 February 1958—I. Abnormal vibration of N_2^+

K. C. Clarke and A. E. Belon: Spectroscopic observations of the great aurora of 10 February 1958—II. Unusual atomic features

M. L. Meeks and J. C. James: Meteor radiant-distributions and the radio-echo rates observed by forward scatter

E. A. Lytle and D. M. Hunten: The ratio of sodium to potassium in the upper atmosphere

J. F. Noxon, A. W. Harrison and A. Vallance Jones: The infrared spectrum of the night airglow from 1·4 μ to 4·0 μ

G. Kvifte: Auroral and nightglow observations at Ås, Norway

H. K. Paetzold: The fluctuations of the acceleration of satellites

R. L. Ingraham: Note on the tidal theory of the S_q magnetic field

J. D. Whitehead: The absorption of short radio waves in the *D*-, *E*- and *F*-regions of the ionosphere

J. A. Fejer and R. W. Vice: An investigation of the ionosphere *D*-region

J. A. Fejer and R. W. Vice: The use of full-wave solutions in the interpretation of ionospheric absorption measurements

J. Egedal: On a pseudo-iterating method for reduction of observations of earth-tide for effects from tides of the sea, with an application to observations made at Bergen (Norway) 1934

K. Bibl, A. Paul and K. Rawer: Die frequenzabhängigkeit der ionosphärischen absorption

D. M. Schlapp: Frequency in the *E*-region of the ionosphere

G. V. Groves: A theory for determining upper-atmosphere winds from radio observations of meteor trails

J. M. Kelso: The determination of the electron density in interplanetary space

VOL. 14 1959

Papers to be published in future issues

J. A. GLEDHILL: The effects of a solar eclipse on a stratified ionosphere

J. A. GLEDHILL: The behaviour of the ionosphere over Cape Town and Johannesburg during the annual solar eclipse

S. R. KHASTGIR: Study of 'winds' in the *F*-region of the ionosphere during the unusual days in the IGY calendar

J. S. GREENHOW and E. L. NEUFELD: Turbulence at altitudes of 80-100 km and its effects on long-duration meteor echoes

H. J. A. CHIVERS and J. S. GREENHOW: Auroral ionization and the absorption and scintillation of radio stars

G. S. KENT: High frequency fading observed on the 40 Mc/s wave radiated from artificial satellite 1957 α

H. J. A. CHIVERS and H. W. WELLS: Observations of unusual radiofrequency noise emission and absorption at 80 Mc/s

W. DIEMINGER: Transient fine structure of the *E*-layer

D. STRANZ: Solar activity and the altitude of the tropopause near the equator

K. C. YEH and O. G. VILLARD JR.: Fading and attenuation of high-frequency radio waves propagated over long paths crossing the auroral, temperate and equatorial zones

O. HACHENBERG and A. KRÜGER: The correlation of bursts of solar radio emission in the cm-range with flares and sudden ionospheric disturbances

M. D. VICKERS: The calculation of the MUF factor for a non-parabolic ionospheric layer

S. DATTA, P. BANYOPADHYAY and R. N. DATTA: Ionospheric observations on the *F*-region during the solar eclipse of April 19, 1958

W. C. HOFFMAN: The ion-electron whistler mode in the earth's dipole field

L. WALLACE: An analysis of a spectrogram of the aurora of February 11, 1958, in the wavelength range 3710–4420 Å

C. M. MINNIS and G. H. BAZZARD: The geographical distribution of ionization in the *F*2-layer

C. M. MINNIS and G. H. BAZZARD: A daily index of solar activity based on *E*-layer ionization (July 1957–December 1958)

D. P. HALE: Characteristics of the visual aurorae at Byrd station, Antarctica, during 1957

B. J. FULTON, L. E. PETRIE and W. S. P. WARD: Transient modes of high-frequency radio wave propagation across the auroral zone

F. H. HIBBERD and J. A. THOMAS: The determination of the electron distribution in the upper ionosphere from satellite Doppler observations

VOL
14
195

VOL.
14
1959